TRAVEL BEHAVIOR DYNAMICS FROM A LONGITUDINAL PERSPECTIVE IN INDONESIA

A Dissertation
Presented to
The Academic Faculty

by

Alyas Abibawa Widita

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of City and Regional Planning

Georgia Institute of Technology
May 2020

TRAVEL BEHAVIOR DYNAMICS
FROM A LONGITUDINAL PERSPECTIVE
IN INDONESIA

Approved by:

Dr. Timothy Welch, Advisor
School of City and Regional Planning
Georgia Institute of Technology

Dr. Brian Stone
School of City and Regional Planning
Georgia Institute of Technology

Dr. Alberto Fuentes
School of International Affairs &
School of City and Regional Planning
Georgia Institute of Technology

Dr. Deden Rukmana
Department of Community and
Regional Planning
Alabama A&M University

Dr. Anjali Thomas
School of International Affairs
Georgia Institute of Technology

Date Approved: March 25, 2020

To hope and resilience

ACKNOWLEDGEMENTS

First and foremost, I am grateful for the opportunity to learn under the tutelage of my advisor, Dr. Timothy Welch. Throughout my academic journey, he has provided invaluable support and encouragement. He has also opened many opportunities for me to work on a multitude of collaborative research projects that have strengthened my capacity to conduct research.

This dissertation work would not have come to fruition without the guidance from my other dissertation committee members: Dr. Alberto Fuentes, Dr. Anjali Thomas, Dr. Brian Stone, and Dr. Deden Rukmana. Their unique insights and perspectives on the substantive topics of this dissertation, as well as suggestions on the details, writings, and formatting, have allowed me to improve this dissertation substantially.

Numerous people have helped me during my doctoral studies at Georgia Tech. I thank all the faculty, staffs, and students at the School of City and Regional Planning; Dr. Catherine Ross and all members of the Center for Quality Growth and Regional Development; Dr. Giovanni Circella, Ali Etezady, and individuals involved in the GSTDM project. I also acknowledge UGM and LPDP for facilitating intellectual growth.

Finally, I cannot thank my wife and son enough for accompanying this journey. I hope it has been a wonderful experience living and growing up in Atlanta. I also thank our extended family for their unconditional love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BHPS	British Household Panel Survey
BMI	Body Mass Index
BPS	*Badan Pusat Statistik*
DID	Difference-in-Differences
DNMP	Dutch National Mobility Panel
EAP	East Asia and the Pacific
GDP	Gross Domestic Product
GSOEP	German Socioeconomic Panel
IDR	Indonesian Rupiahs
IFLS	Indonesian Family Life Survey
INDO-DAPOER	Indonesia Database for Policy and Economic Research
MOP	German Mobility Panel
MXFLS	Mexican Family Life Survey
NFE	Non-Farm Employment
NGO	Non-Governmental Organization
NPU	Neighborhood Planning Unit
NRFOS	National Rural Fixed Observational Site
PMT	Person Miles Traveled
PODES	*Potensi Desa*
PSID	Panel Study of Income Dynamics
PSM	Propensity Score Matching
PSTP	Puget Sound Transportation Panel

TAZ	Traffic Analysis Zone
TOMS	Total Ozone Mapping Spectrometer
TPB	Theory of Planned Behavior
UKHLS	UK Household Longitudinal Study
VMT	Vehicle Miles Traveled
WDI	World Development Indicators

SUMMARY

The notion that the world is rapidly urbanizing is widely known. Indeed, a study has shown that half of the world's population call urban areas their home (United Nations, 2012). The urbanization process occurring in multiple corners around the world can be characterized by three aspects: rural-urban migration, natural growth, and area reclassification from rural to urban (World Bank, 2018a).

This dissertation links this urbanization phenomenon with the literature of travel behavior, particularly as it relates to its interaction with the built environment. The literature has pointed out the relative importance of compact, spatially mixed, and dense environments as a means to shape transportation user behavior. However, the current studies focus mainly on using cross-sectional observation, which limits the collective understanding of the extent of how the dynamic characteristics of the built environment, after controlling for socioeconomic indicators, might exert influences on travel behavior. In addressing this gap, this dissertation proposes three research questions related to two of the three legs of urbanization (World Bank, 2018a), i.e., rural-urban migration and natural growth, to examine the dynamic aspects of travel behavior. Indonesia was selected as the case study considering the rapid urbanization process occurring in the country as well as the presence of the longitudinal Indonesian Family Life Survey (IFLS) data that allows the exploration of the research questions.

The first research question explores travel behavior changes from the lens of rural-urban migration. It finds that relocating to urban areas could reduce household transportation expenditure, as a proxy for travel demand, by approximately 10% relative

to the ones who remained rural. The second research question addresses the natural growth aspect by examining how changes in the built environment influence transportation expenses over time for urban non-movers. It finds the modest, inelastic, and insignificant relationship between gross household density and household transportation expenditure, which also speaks to the relative stability of travel behavior for non-movers even when the physical characteristics of the neighborhood they live in evolve considerably rapidly. The third research question examines which life stage could the built environment influences present walking behavior. It finds that higher exposure to dense environments during childhood could induce a greater likelihood of maintaining walking habits during adulthood.

Collectively, results from these analytical chapters highlight the notion of 'windows of opportunity' (Müggenburg et al., 2015; Prillwitz et al., 2006), where travel behavior might be shaped through life events and past experiences. These findings could have several policy implications. For one, the results indicating that 'sudden' exposure to the denser environment could induce travel behavior changes (i.e., reducing household transportation expenses) suggest the relative importance of ensuring the supply of compact and connected environment, which supports the proposition of building up, e.g., through densification, rather than spreading out or outward growth. Moreover, this proposition also holds relevance in light of the aim of promoting a more sustainable travel pattern through increased walking as continued exposure to dense urban environments during childhood appears to induce walking habits during adulthood.

CHAPTER 1. INTRODUCTION

1.1 Problem Statement

Rapid urbanization is the prominent phenomenon of this century. Countries in East Asia and the Pacific (EAP) region are likely to experience a significant degree of urbanization pressure given its status as "the most rapidly urbanizing region globally" (Baker & Gadgil, 2017, p.5). With an average urbanization rate of 3%, the EAP region urbanized more rapidly than the global average of 2.06% (Baker & Gadgil, 2017).

The urbanization process in the EAP region, as is common in other parts of the world, indicates the increasing share of the region's population living in urban areas and the transformation from a rural to the urban economy (Baker & Gadgil, 2017; World Bank, 2016). Specifically, as outlined by the World Bank (2018a), the urbanization process entails three aspects: rural-urban migration, natural growth, and area reclassification. What makes the urbanization process in most countries in the EAP region somewhat remarkable is how rapid the process has been in comparison to the similar process that occurred in Europe and North America decades ago (European Union, 2016; Sheng, 2012; World Bank, 2015).

The urbanization process and its associated societal and built environment change, therefore, offers a fitting background to study travel behavior dynamics or observation of travel behavior outcomes and its associated factors from more than one specific point in time. Studies have suggested that the built environment – or macroscopic – (e.g., population and employment density) and socioeconomic – or microscopic – (e.g., income, household size) factors are the primary factors influencing travel behavior (e.g., vehicle miles traveled (VMT), transportation expenditure, vehicle ownership) (Ewing & Cervero, 2010; Guerra et al., 2018; Kitamura, 1990; Stevens, 2017). One particular empirical highlight of the literature, and the subject of continued debates, is the collective findings that posit the relative importance of built environment measures to influence travel behavior (Ewing & Cervero, 2010; Stevens, 2017).

However, despite the notion that literature on travel behavior is extensive (Handy, 2017), the current studies estimating the extent of the built environment and socioeconomic factors on travel behavior rely heavily on cross-sectional (Coevering et al., 2015; Næss et al., 2018; Scheiner, 2007). This proposition limits collective understanding of the relative influences of changes in the built environment on travel behavior, after controlling for the dynamic aspect of socioeconomic characteristics as well. For instance, while the literature has generally found the association between higher density and lower VMT (Vehicle Miles Traveled), it remains unclear whether exposure to a new or different neighborhood with higher density could indeed reduce travel demand.

Moreover, in light of the notion that "…(t)he single constant about travel behavior is that it is constantly changing" (Long, 1997, p.xv), travel behavior scholars have lamented the importance of analyzing travel behavior associated with life-course events and built environment, sociodemographic, and policy changes over time (Brathwaite & Walker, 2018; Coevering et al., 2016; Kitamura, 1990; Long, 1997; Paaswell, 1997; Pendyala & Bhat, 2017; Scheiner, 2007; Smart & Klein, 2018b). This proposition emphasizes the needs to explore travel behavior dynamics using the meaningful longitudinal dataset, which to date remain sorely lacking in the literature (Brathwaite & Walker, 2018; Coevering et al., 2016; Long, 1997; Næss et al., 2018; Pendyala & Bhat, 2017).

1.2 Research Questions

Considering the problem statement as described above, this research, therefore, aims to bridge the contemporary urbanization phenomenon, characterized by the three aspects as outlined by the World Bank (2018a), with the gap in travel behavior literature on the lack of longitudinal explorations. In doing so, this dissertation seeks to address the central research question as follows: **How does the evolving built environment and socioeconomic factors, associated with the urbanization process, influence travel behavior, particularly from the perspective of rural-urban migrants, urban non-movers, and childhood experiences?**

The focus on the urbanization process as the background phenomenon of this dissertation stems from the relative importance of urbanization from the lens of development literature. One prevailing view in the literature considers urbanization as one component alongside industrialization, agriculture modernization, and infrastructure provision that could pave the way for developing countries to achieve prosperity (Escobar, 2011).

Embarking from this central research question, considering the relative importance of urbanization for the development trajectory of developing countries like Indonesia, as well as following World Bank's model (2018a) that posits urbanization is partly driven by rural-urban migration and natural growth or densification of the existing urban environment, this dissertation proposes three major empirical research questions:

1. How is a household's travel behavior affected by a move from rural to urban location?
2. For urban non-mover households, how do the dynamic of the built environment and socioeconomic changes influence travel behavior?
3. Do socioeconomic traits and built environment exposure during childhood that forms childhood experiences influence walking behavior during adulthood?

The theoretical framework of mobility biographies offers an appropriate lens to examine these research questions. This theoretical framework considers the relative importance of dynamic aspects to examine travel behavior, particularly as it relates to life-course or events a given individual or household might encounter (Lanzendorf, 2003; Scheiner, 2007).

Case selection. Indonesia offers a compelling case to address the above-mentioned research questions, considering the relatively notable urbanization process the country has experienced. A recent study suggests that Indonesia has among the highest urbanization rate amongst countries in the EAP (East Asia and the Pacific) region, which by itself is "the most rapidly urbanizing region globally" (Baker & Gadgil, 2017, p.5). Specifically, Indonesia's urbanization rate at 4.1% per annum is higher than the EAP region's average at 3% and the global average at 2.06% (Baker & Gadgil, 2017; World Bank, 2016). Figure 1 further corroborates this notion indicating Indonesia's rapid urbanization relative to several neighboring countries and regions across the globe. As can be seen, Indonesia's urbanization rate rose dramatically from 14.6% in 1960 to 50.6% in 2011, which is similar to the phenomenon observed in China. Indeed, both countries only took approximately 50 or more years to increase the share of the respective country's urban population from roughly 10% to 50%. Acceleration of urbanization rate that was not only outpacing the

EAP region's average, but also other regions, i.e., Europe, North America, and Latin America and the Caribbean.

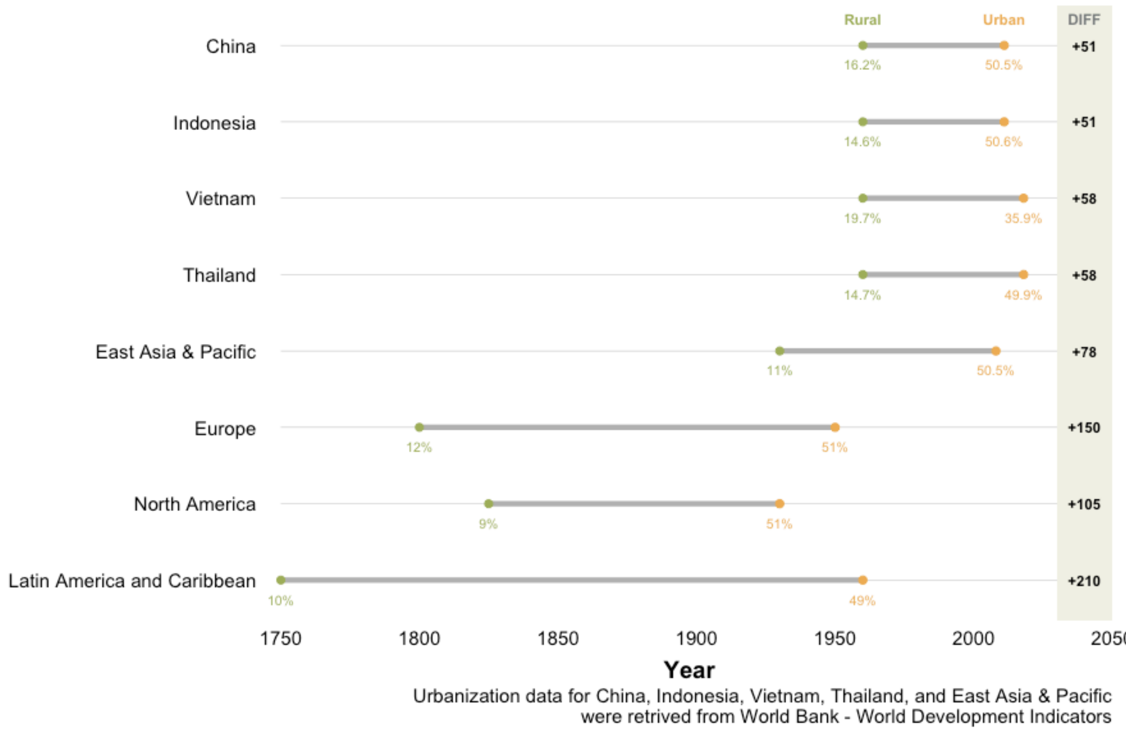

Figure 1 – Indonesia's average urbanization rate in comparison to various countries and regions around the world. Values indicate the share of the countries' and regions' population living in urban areas.

(Source: Asian Development Bank (2012) estimates based on previous studies (Bairoch, 1988; United Nations, 2012); Author's estimates using World Development Indicators (WDI) data (Arel-Bundock, 2019; World Bank, 2018b))

Another factor that substantiates the notable urbanization process in Indonesia is the change in the average density of the country's urbanized areas over the past decade. A 2015 study reveals the substantial increase of average density at 1,974 people/km^2 from 2000 to 2010 in Indonesia, a magnitude that was exceptional in comparison to most countries in the EAP region during the same period (World Bank, 2015).

These factors, as summarized above, imply that Indonesia could have a relative advantage and, therefore, appropriateness for observing travel behavior dynamics in comparison to other countries that had lower urbanization rates. This proposition is particularly relevant considering the theoretical framework of travel behavior that posits the relative importance of the built environment, as discussed further in Chapter 2.

1.3 Summary of Findings

In addressing the first research question exploring travel behavior changes from the lens of rural-urban migration, the analyses revolve around linking two waves of the Indonesian Family Life Survey (IFLS) data, i.e., IFLS 4 (2007) and 5 (2014). Specifically, this dissertation exploits the characteristic of the survey design that tracks individuals and households over time, even if the individuals and households in question relocated to different areas. This somewhat unique characteristic of the survey data allows the application of a research design that revolves around a before-after treatment-control evaluation. That is, relocating households are assigned into the treatment group while those who remained rural are assigned into the control group, which is identified through Propensity Score Matching (PSM) to ensure both treatment and control share similar socioeconomic characteristics. A descriptive analysis using a t-test suggests that relocating households appear to have their household transportation spending reduced quite substantially in comparison to rural households. Subsequent analyses using the Difference-in-Differences (DID) estimation approach indicate, albeit statistically insignificant, that relocating to urban areas could reduce household transportation expenditure, as a proxy for travel demand, by approximately 11% relative to the ones who remained or relocated to the rural environment.

The second research question addresses the natural growth aspect of the urbanization process by examining how changes in the built environment influence transportation expenses over time for urban non-movers. Similar to the first research question, two waves of the IFLS data, i.e., IFLS 4 (2007) and 5 (2014), are used to address the second research question. The analyses also make use of the Indonesian census data (*Potensi Desa* – PODES) that provide the built environment indicators of interest. A descriptive analysis indicates the relative stability of travel behavior for urban non-movers where average household transportation spending does not change substantially between 2007 and 2014, particularly in comparison to rural-urban households as elaborated in the first analytical chapter. Further analyses using panel regression models suggest the inelastic, insignificant, and modest relationship between gross density and household transportation expenditure.

The third research question examines at which life stage, i.e., childhood versus adulthood, could the built environment and socioeconomic factors influence walking behavior during adulthood by pooling an approximately 15-year worth of dataset derived from the IFLS 3 (2000) and 5 (2014). The empirical strategy thus relies on testing the model performance between childhood and adulthood model using mixed-effects logistic regression models to explore the factors associated with walking habits during adulthood. The model comparison indicates that the childhood-only model offers a better goodness-of-fit than the adulthood-only model. One particular built environment indicators of interest suggest that greater exposure to dense environment during childhood could induce a walking habit during adulthood.

Collectively, results from these analytical chapters highlight the notion of 'windows of opportunity' (Müggenburg et al., 2015; Prillwitz et al., 2006), where travel behavior might be shaped through life events, specifically rural-urban migration, and experiences during childhood.

1.4 Outline of the Dissertation

Following the first chapter of Introduction, this dissertation is structured in an interconnected manner as follows. In Chapter 2, the emphasis is given to describe the theoretical underpinnings of travel behavior, particularly as it relates to the built environment, present empirical studies and the prevalence of cross-sectional data, and the potential application of a panel survey for exploring the dynamic aspects of travel behavior.

Chapter 3 elaborates on the case selection (i.e., Indonesia), the overall research framework, and the longitudinal Indonesian Family Life Survey (IFLS) as the primary data source. It also discusses how the data, notably the dependent variables (i.e., household transportation expenditure and binary indicator of walking habit), corresponds with the present travel behavior literature.

The subsequent analytical chapters from Chapter 4 to 6 address each of the three research questions: 1) rural-urban migration, 2) panel analyses of urban non-movers, and 3) past experiences. Each of the analytical chapters presents in further detail the literature related to the substantive topic, the research framework, the methodologies applied, and the results as well as contributions to the literature.

Finally, Chapter 7 summarizes and links the results and presents potential research and policy implications drawn collectively from the findings.

CHAPTER 2. LITERATURE REVIEW

The literature review chapter is structured as follows: The first section will discuss the state of travel behavior literature, mainly as it revolves around the relationship between travel behavior and the built environment. Emphases are explicitly given to discuss the theoretical framework, current debate, and geographical coverage of the studies within this strand of literature. In the second section, an elaborative description of the mobility biographies approach that acts as the theoretical guidance of this dissertation is presented. The mobility biographies approach emphasizes observing travel behavior outcomes as a function of life-course trajectory, which lends a hand to the importance of longitudinal data. To this end, in the third and final section of this chapter, the focus is to discuss the extent of longitudinal data application in the travel behavior literature.

2.1 Travel Behavior

2.1.1 Theoretical Framework

The theoretical foundation of travel behavior stems from the notion that "transportation is a derived demand," or the "desire to undertake varying activities" (Small & Winston, 1999, p. 48). That is, people rarely travel or its own sake, but rather as a means to conduct or participate in particular activities. In Boarnet and Crane's (2001, p. 65) words, "people typically travel as a means to an end, not as an end in itself."

Given this derived-demand perspective, researchers mostly consider travel as a disutility due to, in parts, the time lost in travel (Kraus, 1977). Most transportation researchers, therefore, subscribe to the widely accepted assumption and paradigm that people are seeking to maximize their travel utility (or minimize their travel disutility).

Analyses of transportation users' behavior and most travel demand models, both aggregate and disaggregate, are primarily grounded on this utility-maximization principle. As a multitude number of transportation users' behavior analyses has shown, it appears predominant determinants influencing travel behavior can be grouped into the socioeconomic and built environment factors.

In terms of socioeconomic factors, studies have shown the extent to which variables under this factor influence travel behavior. For instance, the demographic factor is an essential factor since transportation users likely have different values that they consider regarding the characteristics and "quality attributes" of specific transportation options (Small & Winston, 1999, p. 12). In a study by Lave (1969) of 280 commuters in Chicago, he found that female commuters are more likely to take transit than male commuters, all else equal. Moreover, the economic factor appears to be substantially profound in influencing travel behavior, particularly income. As income increases, households or individuals could afford a more variety of transportation options and eventually chose the alternative that could subjectively offer the highest utility, which in many cases tends to be the private vehicle. Other socioeconomic factors, such as vehicle ownership, household size, the presence of child(ren), are also regularly incorporated to model travel behavior; however, socioeconomic covariates are often considered as control variables rather than the primary policy variables to be evaluated.

The built environment factor, which often acts as the policy variables of interest, relates to the notion of transportation as a means to reach "varying activities distributed over time and space" (Small & Winston, 1999, p. 48). That is, the built environment factor as represented by how activities are spatially distributed at a given place, city, or the broader region would theoretically influence travel utility and, therefore, transportation users' behavior. Specifically, built environment factors likely influences travel behavior through the function of "how far individuals are from destinations, and it is the cost of this distance that influences where they can go, by what mode, and how frequently" (Handy, 2017, p. 26). One could imagine the conventional logic of, say, policy on increasing land use mix in a given neighborhood to concentrate activity centers and, therefore, may reduce the length of auto trips.

Incorporating demographic, economic, and built environment factor as described above, Boarnet and Crane (2001) proposed three formal models to represent the general modeling approach to estimate built environment (**L**) effects on travel behavior (*a*), controlling for income (y) and socio-demographics (S):

I. $a = f(\mathbf{L}, \mathrm{y}, \mathrm{S})$

II. $a = f(p, \mathbf{L}, \mathrm{y}, \mathrm{S})$

where *p* is trip cost-related variables, e.g., Zegras (2004) used average walk trip time between each zone as a proxy for individual walk trip costs for estimating the number of discretionary walking trips in Santiago de Chile.

III. Two-step procedure: 1) $p = f(\mathbf{L})$ and 2) $a = f(p_e, \mathrm{y}, \mathrm{S})$

where step 1 is used to estimate the effects of the built environment on trip prices (*p*) and step 2 incorporates the effects of predicted prices (p_e) on travel behavior (*a*)

In sum, researchers commonly subscribe to the notion of transportation as a derived demand in exploring factors associated with travel behavior. This demand-driven perspective leads to the widely accepted assumption that travel is a disutility and that transportation users will seek to mitigate their disutility by maximizing utility. Over the years, researchers have discovered to what extent do socioeconomic and built environment factors contribute to influencing travel behavior.

2.1.2 Current Debate

While researchers have identified the demographic, economic, and built environment factors influencing travel behavior, the current debate largely stems from the magnitude of built environment effects on travel behavior. The allure of built environment factor to shape travel behavior stems from the idea that, unlike socioeconomic determinants, the built environment can be shaped through local, municipal, and even national policies and is well situated within the domain that planners, engineers, policymakers, and community could relate to (Boarnet & Crane, 2001; Ewing & Cervero, 2010; Stevens, 2017).

While the appeal of the built environment as a policy tool to shape travel seems understandable, a multitude number of empirical studies have shown that built environment effects on travel appear to be relatively modest. These findings have ignited a persistent debate on the efficacy of the built environment to mitigate the varying consequences of excessive travel. Most recently, a debate appears in the Journal of the American Planning Association (Volume 83, Issue 1), where several researchers responded to Stevens's (2017) article titled '*Does Compact Development Make People Drive Less?*'. Stevens's (2017) central thesis revolves around how modest, and weak the compact development effect is on travel behavior, i.e., the elasticity of household/population density on VMT was -0.22/-0.10. Based on his findings, he then suggests that "[P]lanners and municipal decision makers should not rely on compact development as their only strategy for reducing driving unless their goals for reduced driving are very modest and can be achieved at a low cost" (Stevens, 2017, p. 8).

In response to Stevens's (2017) argument, Ewing and Cervero (2017), Knaap, Avin, and Fang (2017), and Handy (2017) argue the potential travel behavior impacts could be substantial depending upon the context and background of policy initiatives. Knaap et al. (2017, p. 36) put forward a scenario of increasing density in a given area from "2,000 persons or 800 homes per square mile" to "2,500 persons or 1,000 homes per square mile – just 200 more residential units…" could substantially decrease VMT "by 550 miles for each of the 2,500 residents," which is not a measly reduction at all. Along the same line, Handy (2017, p. 26) notes the potential 9% VMT reduction as the result of "40% increase in density, achieved by 30 U.S. cities in a 40 year period" would offer "a good share of the reduction that California needs."

This long-standing debate has mostly been driven and informed by inferences from cross-sectional observations. It remains unclear how the arguments, as mentioned above, apply to a novel research design involving longitudinal assessment. For instance, if higher density could reduce the distance traveled, would relocating to places with higher density could indeed exert that distance-reduction effect? Also, for non-movers, would an increase in density associated with the natural growth of the neighborhood they live in could also exert the said effect?

In addition to the current and long-standing debate as summarized above, several criticisms have been proposed regarding how present studies approach the built environment and transportation interaction. One of the long-standing criticisms is the potential multicollinearity among built environment variables included in the model. As expected, denser neighborhoods tend to have greater land use mix and a higher number of four-way intersections, which likely provide a pleasant walking experience but a rather unpleasant driving experience (Manville, 2017). To address this issue, researchers typically conduct a statistical test to diagnose potential collinearity and remove a particular variable(s) based on a specified threshold.

Another issue is related to spatial autocorrelation that could arise where researchers typically aggregately measure built environment attributes based on a given spatial unit, e.g., traffic analysis zone (TAZ), block groups, or census tracts (Hong et al., 2014; Zhang & Zhang, 2018). Assigning these attributes to the individual or household level data might threaten "the assumption of independence among observations for statistical analysis, which produces statistically inefficient coefficients and likely leads to erroneous conclusions" (Zhang & Zhang, 2018, p. 5). Researchers recognize this issue as an *ecological fallacy* (Holt et al., 1996; Robinson, 1950).

One of the potential remedies to address this ecological fallacy problem is by using specified individual or household's longitude and latitude coordinate data and build the built environment attributes around this highly-disaggregated information (Ewing et al., 2013). However, this approach might be impractical since most spatial data sources are aggregated already at a given spatial unit and the possible privacy concerns regarding access to the x and y coordinates of study respondents. Moreover, researchers might instead fall into the opposite ends of the ecological fallacy problem, that is, the *atomistic fallacy* or the problem of overlooking "the contextual effects at the aggregate level" (Zhang & Zhang, 2018, p. 5).

To alleviate the concerns over ecological and atomistic fallacy, researchers reflected upon the underlying problem that motivated the research. In the case of travel and built environment research, researchers could use the spatial units where the policy intervention is evaluated. For instance, if the aim of the research was to assess the how built environment changes at the Neighborhood Planning Unit (NPU) could impact travel behavior, then the NPU should consistently be used to develop the built environment attributes (Zhang & Kukadia, 2005; Zhang & Zhang, 2018).

2.1.3 Data Sources and Geographical Coverage

2.1.3.1 <u>Data Sources</u>

A survey of the literature indicates that most travel behavior studies rely on household travel survey data. Researchers typically use this data collected by regional transportation agencies in a given city or region, with a few exceptions where the researchers collected the data on their own to better suit particular research objectives. This notion is reflected in the list of several studies, as summarized in the meta-analysis papers by Ewing and Cervero (2010) and, more recently, Stevens (2017), as shown in Table 1. The decision to focus on these two studies stems from the notion that both are considered as the well-cited, comprehensive, and relatively recent sources summarizing the travel behavior and built environment literature. However, not every study reviewed in both meta-analysis papers is discussed to avoid overrepresenting certain regions that have been heavily analyzed over the years, e.g., San Francisco Bay Area, CA, Portland, OR, to name a few. Moreover, it should also be noted that several studies reviewed in these two articles are overlapped.

Table 1 – Articles on travel behavior and built environment interaction reviewed in meta-analysis papers

Study	Data	Source	Type
Asad (2013)	Scottish Household Survey	GA	CS
Bento, Cropper, Mobarak, and Vinha (2003)	Nationwide Personal Transportation Survey	GA	CS
Bhat, Sen, and Eluru (2009)	Bay Area Household Travel Survey	GA	CS
Boarnet, Greenwald, and McMillan (2008)	Portland Metro Travel Diary	GA	CS
Boarnet, Joh, Siembab, Fulton, and Nguyen (2011)	South Bay Area, Los Angeles Travel Survey	OR	CS
Boer, Zheng, Overton, Ridgeway, and Cohen (2007)	Nationwide Personal Transportation Survey	GA	CS
Cao, Handy, and Mokhtarian (2006)	Austin, TX Travel Survey	OR	CS
Cao, Mokhtarian, and Handy (2009b)	Northern California Travel Survey	OR	CS
Cervero (2002)	Metropolitan Washington Council of Government (MWCOG) Travel Survey	GA	CS
Cervero and Duncan (2006)	Bay Area Household Travel Survey	GA	CS
Cervero and Kockelman (1997)	Bay Area Household Travel Survey	GA	CS
Chatman (2009)	San Francisco, CA/San Diego, CA	OR	CS
Fan (2007)	Greater Triangle, North Carolina Travel Survey	GA	CS
Frank, Bradley, Kavage, Chapman, and Lawton (2008)	Puget Sound Regional Council (PSRC) Household Travel Survey	GA	CS
Guerra (2014)	Mexico City Metro Travel Survey	GA	CS
Handy, Cao, and Mokhtarian (2006)	Northern California Travel Survey	OR	CS

Study	Data	Source	Type
Heres-Del-Valle and Niemeier (2011)	California Statewide Household Travel Survey	GA	CS
Holtzclaw, Clean, Dittmar, Goldstein, and Haas (2002)	Chicago Area Transportation Study/Southern California Association of Government (SCAG)/Bay Area Metropolitan Transportation Commission (MTC)	GA	CS
Khattak and Rodriguez (2005)	Chapel Hill and Carrboro, NC Travel Survey	OR	CS
Kitamura, Mokhtarian, and Laidet (1997)	San Francisco Bay Area Travel Survey	OR	CS
Kuzmyak, Baber, and Savory (2006)	Baltimore Metropolitan Council (BMC) Travel Survey	GA	CS
Lee and Moudon (2006)	Seattle Metro/King County Travel Survey	OR	CS
Majid, Nordin, and Medugu (2014)	Iskandar Regional Development Travel Diary	OR	CS
Nasri and Zhang (2012)	Washington, D.C.; Baltimore, MD; Seattle, WA; Atlanta, GA; Richmond–Petersburg, VA; Norfolk– Virginia Beach, VA	GA	CS
Rodriguez and Joo (2004)	UNC Chapel Hill Commuter Survey	OR	CS
Zahabi, Miranda-Moreno, Patterson, and Barla (2015)	Montreal, Canada Origin-Destination Survey	GA	CS
Zegras (2010)	Santiago de Chile Household Origin-Destination Survey	GA	CS
Zhang (2004)	Boston, MA and Hong Kong	GA	CS
Zhou and Kockelman (2008)	Austin, TX Area Household Travel Survey	GA	CS

Note:

Source of data	GA	: Government agencies
	OR	: Original data collected by authors
Type of data	CS	: Cross-sectional

Considering that the list of travel behavior studies, as shown in Table 1, represents the state of the literature, it might be inferred that virtually every empirical travel behavior study relies on cross-sectional data. Explorations of travel behavior using longitudinal data remain curiously lacking in the literature.

2.1.3.2 <u>Geographical Coverage</u>

The list, as shown in Table 1, also provides a glimpse of the geographic distribution of empirical studies in the travel and built environment literature. As can be seen, it is apparent that the predominant share of the studies originated from cases in the U.S. An investigation of the geographical coverage of travel and built environment literature was conducted by exploring the distribution of empirical cases as appeared in the Journal of Transport and Land Use[1]. The focus on this platform stems from the substantive theme that this journal explicitly addresses, i.e., the travel and built environment interaction.

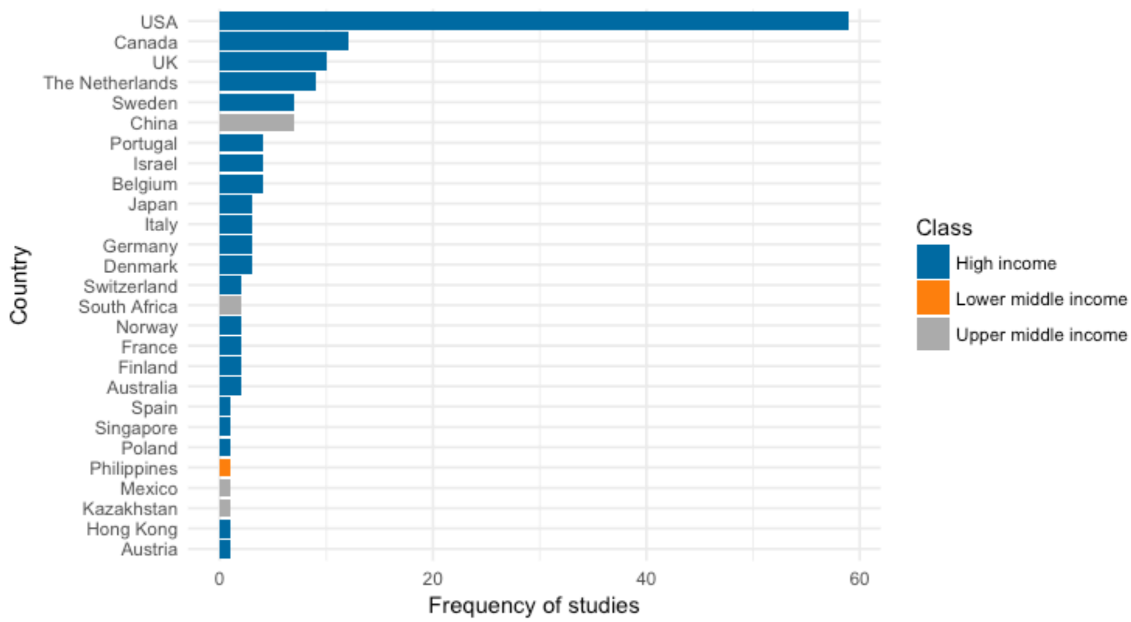

Figure 2 – Geographical distribution of 148 empirical cases in the Journal of Transport and Land Use, 2008 – 2017

Figure 2 further corroborates the notion as indicated in Ewing and Cervero's (2010) and, more recently, Stevens's (2017) meta-analysis papers that travel behavior and built environment scholarship has been noticeably focused in the U.S. and several other high-income countries. Explorations of empirical cases within the travel behavior literature

[1] Journal homepage: https://jtlu.org/index.php/jtlu

under different contexts, particularly from lower middle-income countries, might very well be necessary to provide researchers, planners, and policymakers with a more comprehensive picture of comparative travel behavior analyses and generalized understandings of the dynamics across the globe.

2.2 Mobility Biographies

2.2.1 Theoretical Framework

It is apparent from the review of the literature, as discussed in the previous sections, that the prevailing studies in travel behavior and built environment scholarship overly rely on cross-sectional analyses. Considering that this dissertation focuses on the exploration of travel behavior dynamics, the theoretical framework of "mobility biographies" is applied to guide the overall research endeavor.

The mobility biographies concept stems from the life-course approach (Chatterjee & Scheiner, 2015; Lanzendorf, 2003; Müggenburg et al., 2015; Scheiner, 2007). One of the pioneering studies on the life-course approach was the research pioneered by W.I. Thomas and Florian Znaniecki, who studied Polish peasants and how they characterized themselves as they were growing older and got exposed to a variety of societal changes (Thigpen, 2017). Building off of the life-course approach as pioneered in the early 20th century, a variety of propositions on the definition of a life-course approach have been proposed. The basic idea of the life-course approach, as Chatterjee and Scheiner (2015, p. 5) suggested, "is a multidisciplinary paradigm for the study of people's lives, structural contexts and social change."

The emphasis on the temporal factor that the life-course approach is built upon, therefore, seems fitting with the argument that "…[T]he single constant about travel behavior is that it is constantly changing" (Long, 1997, p. xv). The mobility biographies approach takes into account this apparent linkage by proposing a set of life-course factors or domains that might be interrelated with travel behavior (Chatterjee & Scheiner, 2015; Lanzendorf, 2003; Müggenburg et al., 2015; Scheiner, 2007). As a note, while the term of mobility biographies has arguably become the formal vocabulary representing how life-course approach is incorporated into travel behavior studies, scholars have to some extent touched upon this subject several decades ago. This is particularly apparent in, for instance, Salomon and Ben-Akiva's (1983) paper on the influence of lifestyle choices on travel demand (Chatterjee & Scheiner, 2015; Lanzendorf, 2003).

However, it was not until in the early 2000s that scholars have formally constructed how to incorporate the life-course perspective in travel behavior research – marking the advent of mobility biographies as recognized today. Two primary references, or canonical papers, on mobility biographies are the works by Lanzendorf (2003) and Scheiner (2007) (Müggenburg et al., 2015).

In his paper titled *Mobility biographies: A new perspective for understanding travel behavior*, Lanzendorf (2003) suggests three primary interrelated life domains that are thought would influence travel behavior: *lifestyle domain*, *accessibility domain,* and *mobility domain*. Lanzendorf's (2003) proposition extends the framework as proposed by Salomon and Ben-Akiva (1983) by adding the temporal considerations (Chatterjee & Scheiner, 2015; Müggenburg et al., 2015). The mobility biographies framework that Lanzendorf (2003) proposed considers the habitual or stability aspect of travel behavior.

Changes in and the dynamics of travel behavior are, therefore, likely attributed to the life-course of a given individual or household following the three domains as indicated (Lanzendorf, 2003). The *lifestyle domain* mainly refers to demographic characteristics (marriage, household composition) and participation in the workforce. *Accessibility domain* includes residential location, employment location, and overall urban form characteristics. The *mobility domain* refers to access to cars, transit, or other means of transportation, as well as the daily travel behavior itself.

Along a somewhat similar line as Lanzendorf (2003), Scheiner (2007) proposes the interrelationship between mobility biography and several related ‘partial biographies’ (i.e., employment biography, household biography, and residential biography). Scheiner (2007) argues that the approach to understanding *mobility biography* should not overlook these relevant ‘partial biographies’ while recognizing that each biography could exert different effects on travel behavior. For instance, Scheiner (2007) notes that one example of *residential biography* like changes in the characteristics of the physical environment surrounding a given individual’s residential location likely takes place gradually; hence its effects on short term travel behavior would likely be marginal. Certain moments in other biographies such as the formation of a household under *household biography* or retirement under *employment biography* might exert stronger influences on travel behavior than gradual and marginal changes in residential location for non-movers (Scheiner, 2007). However, Scheiner (2007) suggests that one aspect of residential biography that is likely to exert a noticeable influence on travel behavior is residential relocation.

Table 2 below indicates the apparent overlap between Lanzendorf’s (2003) and Scheiner’s (2007) framework. As indicated, the *accessibility domain* that Lanzendorf

(2003) proposed shares more or less similar notions as the *residential biography* in Scheiner's (2007) framework. Intuitively, Scheiner's (2007) *household* and *employment biography* comprised of components that overlap with Lanzendorf's (2003) *lifestyle domain*. One aspect that is being highlighted in Scheiner's (2007) paper but not in Lanzendorf's (2003) is the likely influence of childhood experiences. Both authors subscribe to the hypothesis that travel behavior is a reasonably stable habit. However, there seems to be no obvious emphasis on past experiences, or childhood specifically, in Lanzendorf's (2003) paper.

Table 2 – Comparison of the mobility biographies framework

	Paper	
	Lanzendorf (2003)	**Scheiner (2007)**
Element	Accessibility domain	Residential biography
	Lifestyle domain	Household biography
		Employment biography
	Mobility domain	Mobility biography

Guided by the theoretical framework of mobility biography, this dissertation and the research questions address the elements of and interrelations between mobility biography and the related biographies as outlined above. The first research question focuses on residential relocation while simultaneously incorporates indicators of household biography, employment biography, and mobility biography. Along a similar line, the second research question considers the aforementioned biographies while focusing on non-movers. The third research question explores how relevant biographies during childhood and adulthood influence current travel behavior.

2.2.2 Methodological Approaches of Mobility Biographies

Several methodological considerations have been identified to conduct travel behavior research under the theme of mobility biographies. As elaborated by Lanzendorf (2003) and Scheiner (2007), there are three principal data collection methods that would allow analyses of travel behavior dynamics under the mobility biographies approach: panel survey, retrospective survey, and pseudo-panels from repeated cross-sectional observations. Chatterjee and Scheiner (2015) extend the methodological considerations by adding the potential of life-history interviews. This qualitative approach that might allow the respondents to address open-ended questions as a means to examine how travel behavior relates to their life-course.

Panel survey. Panel survey appears to be the ideal option as it could capture "comprehensive mobility biographies" over a given individual's or household's life-course; hence travel behavior dynamics can be examined (Scheiner, 2007, p. 170). While panel survey offer advantages as it could reveal the dynamics surrounding travel behavior, conducting the studies using such approach is often requires substantial efforts and time (Lanzendorf, 2003; Müggenburg et al., 2015; Scheiner, 2007).

In recent years, however, scholars have started to notice the opportunity provided by a somewhat unusual source of data to explore travel behavior dynamics; that is, through the application of the household panel survey. Household panel survey might not offer a comprehensive set of travel behavior information such as what travel mode was being used, for how long, what was the trip purpose, to name a few. Nonetheless, a multitude of household panel surveys provides transportation-related information (e.g., vehicle ownership, transportation expenditure) that could be used to conduct transportation research.

Retrospective survey. While the panel survey is the ideal option, most of the time, it might not be the feasible option (Lanzendorf, 2003; Scheiner, 2007). As indicated in greater depth in the literature review section below, several researchers have adopted a retrospective survey or probing information where the respondents were asked to recall past experiences and relate them with travel behavior outcomes. However, some have cast doubts on the validity of data from retrospective surveys as respondents might not be able to recall their past experiences or key events accurately, thus might undermine the validity of the analyses (Lanzendorf, 2003; Müggenburg et al., 2015; Scheiner, 2007).

Pseudo-panel survey. In comparison to the panel survey and retrospective survey, the pseudo-panels approach looks to be the least ideal option (Lanzendorf, 2003; Scheiner, 2007). While the abundance of repeated cross-sectional data would make it convenient to construct pseudo-panels, this approach aggregates the observations into a given cohort or arbitrary geographic scale. This apparent disadvantage would not allow observations at the individual level and hence fails to satisfy "a genuine biographical approach" (Lanzendorf, 2003; Scheiner, 2007, p. 170). Besides, there are concerns over the suitability of pseudo-panels since the dataset was constructed from long-term time-series cross-sectional data that typically was not tailored to address specific or relevant research questions under the theme of mobility biographies research (Lanzendorf, 2003).

Considering an overview of mobility biographies and the methodological considerations as outlined above, the following sections will further elaborate on the extent of longitudinal analyses in travel behavior literature. Particularly as it relates to mobility biographies and the application of true panel data, followed by more elaborate descriptions of the application of the household panel survey to conduct transportation-related research.

2.3 Panel Data

As noted in the previous section, the panel survey is an ideal source and particularly more suited to observe travel behavior dynamics than data derived from retrospective or pseudo-panel survey. These following sections further elaborate on the relative advantages and disadvantages of panel data followed by an overview of the gradually growing number of travel behavior studies using panel data, particularly household panel survey, that fall under the relevant research theme of mobility biographies.

2.3.1 Advantages and Disadvantages of Panel Data

2.3.1.1 <u>Advantages of Panel Data</u>

In contrast to cross-sectional data focusing solely on variation 'between' respondents in a given time, the application of longitudinal analysis offers a lens to observe travel behavior dynamics 'within' respondents over time (Kroesen & Goulias, 2016). Given this nature of panel data, researchers have argued that the panel data is, therefore, more appropriate than the cross-sectional ones to conduct causal analysis, evaluate (transportation) policy impacts, estimate the effects of particular events, and uncover behavioral changes (Coevering et al., 2016; Kitamura, 1990; Kroesen & Goulias, 2016; Raimond & Hensher, 1997). Indeed, the primary advantage of using panel data is its potential to institute robust causal inference as noted by Raimond and Hensher (1997, p.

16), "The endearing feature of a panel is its ability to capture behavioral changes over time in a level of detail that is necessary to unravel true causality." This particular advantage of panel data is further augmented by the relevance of this type of data for travel behavior research considering the notion that "…[T]he single constant about travel behavior is that it is constantly changing" (Long, 1997, p. xv) and given the theoretical proposition of mobility biographies where travel behavior is shaped by the life-course of a given individual or household (Lanzendorf, 2003; Scheiner, 2007).

Another advantage of panel data from the procedural and logistical point of view is that implementing a panel survey might instead be more cost-effective than cross-sectional data collection (Raimond & Hensher, 1997; Tourangeau et al., 1997). While this notion seems counterintuitive, a panel survey might enable cost savings since the survey administrators might not need to conduct a lengthy pre-survey sampling process as well as recruiting new respondents that would otherwise be necessary for the cross-sectional data collection (Raimond & Hensher, 1997; Tourangeau et al., 1997). While the rate of cost savings potential may vary, Tourangeau et al. (1997, p. 8) noted that "the costs of re-interviewing a panel maybe 20 to 80 percent less than the costs of obtaining the same information from a new sample."

2.3.1.2 Disadvantages of Panel Data

The disadvantages of panel data largely stem from the relative difficulties to keep the same respondents to participate in the survey over time, or best known as attrition issues. Indeed, attrition is known as the "Achilles heel" of panel data collection (Thomas et al., 2001, p. 559, 2012, p. 109). In addition to the attrition issue, this following list elaborates some of the disadvantages and difficulties associated with panel data:

- Possible higher likelihood of initial non-response rate since respondents were asked to participate in more than one data collection process (Kitamura, 1990).
- Locating respondents, which might be even more troublesome and costly in a multi-wave panel survey (Kitamura, 1990; Thomas et al., 2012).
- "Fatigue effects" or measurement errors associated with the potential decline in the precision of survey reporting (Van Wissen & Meurs, 1989, p. 109).
- "Population representativeness" or the issue with maintaining a panel sample that remains representative to the population of interest (Kitamura, 1990, p. 411). A number of transportation panel surveys have aimed to address this issue using a refreshed or rotating sampling strategy, as shown in the case of the Dutch National Mobility Panel (DNMP), Puget Sound Transportation Panel (PSTP), and German Mobility Panel (abbreviated as MOP).

Recognizing the advantages and disadvantages of panel data as outlined above, it could be inferred that panel data is not a panacea. Nonetheless, given the somewhat limited number of travel behavior studies utilizing panel observations and due to the proposition that investigation of travel behavior dynamics might be most appropriately be done using panel data, the need for shedding light on travel behavior dynamics through panel data investigation is of critical importance (Kitamura, 1990; Long, 1997). The following section discusses a review of transportation-related studies using panel data to illustrate the extent of the literature and methodological considerations.

2.3.2 Transportation Panel Survey

Application of panel data collection in transportation planning and research can be divided into three types of data collection. First, a general-purpose transportation survey where the same respondents were asked to participate in a typical travel survey data collection over time (e.g., Dutch National Mobility Panel and Puget Sound Transportation Panel). Second, project-specific transportation panel survey that usually aims to evaluate specific transportation policies or projects. And third, another form of panel data in transportation domain is a short panel where repeated surveys are conducted consecutively in a particular period of days, known as "the analysis of multi-day travel behavior" (Kitamura, 1990, p. 402) or "short survey panels" (Comendador & López-Lambas, 2016, p. 249). The aim of studies utilizing a multi-day travel survey is typically to observe activity-travel scheduling and stability and regularity of travel behavior (Kitamura, 1990; Pas, 1988). Since the project-specific transportation panel and short panel hold little relevance to this dissertation, a review of studies on these subjects is not discussed in this dissertation.

In terms of the general-purpose transportation survey, the extent of the literature review indicates that there are only a few panel surveys available across the world, as shown in Table 3. Currently, there are two longest-running transportation panel surveys that employed true panel approach of following the respondents from the initial wave to the last iteration (Van Wissen & Meurs, 1989): 1) the Dutch National Mobility Panel (DNMP) (10 waves, 1984-1989) and 2) Puget Sound Transportation Panel (7 waves, 1989-1993). The German Mobility Panel (abbreviated as MOP) is probably the longest-running active transportation panel survey at this moment; however, the MOP employed a rotating panel approach, where the respondents were interviewed on a yearly basis during the three years period and then were replaced with new samples (Zumkeller & Chlond, 2009). Another example of a general-purpose transportation panel survey was the Bay Area Household Panel, which was initiated in 1990, surveying 10,900 households but could not manage to conduct the subsequent wave owing to the lack of funding (Purvis, 1997).

Table 3 – Regional-level general-purpose transportation panel survey, sorted by years initiated

Survey	Initiated	Waves	Baseline Sample Sizes
Dutch National Mobility Panel	1984	12 (1984-1989)	1,764
Puget Sound Transportation Panel	1989	7 (1989-1993)	1,713
Bay Area Household Panel[1]	1990	1 (1990-)	10,900
German Mobility Panel[2]	1994, 1999	1994-2016	-

[1] Initiated but couldn't manage to conduct the second wave of the survey

[2] German Mobility Panel (MOP) was initiated in 1994 covering the West German federal states and was later expanded in 1999 to cover all Germany federal states. Since the MOP is refreshed every three years, the baseline sample sizes in the table are intentionally left blank.

Regarding the attrition issue, experiences from the DNMP and PSTP suggest that this challenge is very much relevant. Neither surveys were able to maintain a substantial retention rate from the baseline sample. The case of the DNMP was particularly telling to indicate the attrition challenge as the survey was only able to maintain a 58% retention rate in its second wave, which was conducted only a year after the first survey (Tourangeau et al., 1997).

2.3.3 Household Panel Survey

A review of the literature indicates that the lack of transportation-specific panel data has somewhat hindered the exploration of travel behavior under the research theme of mobility biographies. In light of this notion, in recent years, researchers have started to leverage the availability of general-purpose household panel survey to study the dynamic aspects of travel behavior. This section discusses the potential application of household panel surveys to conduct transportation-related research followed by an exploration of the current studies.

2.3.3.1 Overview of Household Panel Survey

The general-purpose household panel surveys are typically a multi-purpose survey aimed towards understanding the livelihood dynamics comprehensively over time, covering a wide range of aspects, e.g., socioeconomic, household composition, migration, consumption and expenses, among others. As expected, the multi-purpose nature of these household panel surveys indicates that they were not designed specifically and exclusively for transportation research (Dargay & Hanly, 2003; Hanly & Dargay, 2000); however, several modules within these surveys might be used appropriately to conduct mobility biographies-based research.

There are various forms of household panel surveys fielded in several countries, as shown in Table 4. A multitude number of surveys were initiated and are currently maintained by the government in the respective country, for instance, National Rural Fixed Observational Site (NRFOS) in China and British Household Panel Survey (BHPS). Most of the other surveys are maintained by non-governmental organizations (NGOs) or research and higher-education institutions, e.g., Indonesian Family Life Survey (IFLS), Mexican

Family Life Survey (MXFLS), and Panel Study of Income Dynamics (PSID). A literature review paper by Hao, Wang, and Xie (2014) describes a thorough description of a variety of household panel surveys around the world.

Table 4 – Household panel surveys around the world

Survey	Abbr.	Initiated	Waves
Panel Study of Income Dynamics - USA	PSID	1968	*Multiple*[1]
Malaysian Family Life Survey	MFLS	1976	1976, 1988
National Rural Fixed Observational Site - China	NRFOS	1984	*Multiple*[1]
Nang Rong Survey - Thailand	-	1984	1984, 1994, 2000
British Household Panel Survey	BHPS	1991	*Multiple*[1]
Indonesian Family Life Survey	IFLS	1993	1993, 1997, 2000, 2007, 2014
Chitwan Valley Family Study - Nepal	CVFS	1996	1996 - 2007
Mexican Family Life Survey	MXFLS	2002	2002, 2005, 2009
India Human Development Survey	IHDS	2004	2004, 2011

2.3.3.2 Transportation Research using Household Panel Survey

The multi-purpose nature of the typical household panel survey suggests that this type of survey is not exclusively designed for transportation analyses (Dargay & Hanly, 2003; Hanly & Dargay, 2000). Indeed, these panel surveys were initiated mainly by economists, political scientists, and demographers, not by transportation planners. Nonetheless, in recent years there is a growing list of transportation planning literature using household panel surveys. These studies took advantage of the longitudinal and long-term nature of this survey to explore a variety of transportation-related topics and address research questions that otherwise might be difficult to tease out using cross-sectional observations.

For instance, in a recent study, Ralph (2018) used the PSID data to estimate the influences of car access during childhood on education, employment, and earnings during adulthood. She specified the treatment group as respondents who grew up in a household that lacked access to the car and the control group as respondents who grew up in a household that always owned a car. Her study reveals that growing up in a carless household reduces the likelihood of a given individual to complete education, being fully employed, and obtained more earnings relative to a comparable individual who always had access to a car during the early stages of his/her life.

Along a somewhat similar line of observing past experiences, Smart and Klein (2018b) also used the PSID data to study the role of previous experiences on individuals' inclination to use transit in the later stages of life. Their study indicates that individuals who had past transit exposure are more likely to use transit even if they moved to a new residential location with fewer transit options. They argue that planners and policymakers should move beyond the conventional cost-benefit analysis in evaluating planning projects since its influences could have long-term consequences.

Several researchers have used the British Household Panel Survey (BHPS) to study travel behavior dynamics under the mobility biographies perspective. Hanly and Dargay (2000) modeled the factors associated with the household car ownership from approximately 4,000 households in Waves 3 through 6 of the BHPS data. Their study reveals a more robust correlational estimation from the panel data to shed light on the relative influences of income, household composition (e.g., number of employed adults, pensioner), and locational characteristics on household car ownership (Hanly & Dargay, 2000). A 2003 study also used BHPS data to estimate a variety of indicators that influencing the dynamics of travel time and mode to work (Dargay & Hanly, 2003). Their analyses reveal that individual characteristics and socioeconomic traits are strong predictors of changes in travel time and mode.

A brief overview of the literature, as presented above, suggests the potential application of household panel surveys to conduct transportation research. This proposition also speaks to the strategy that researchers have implemented to address research questions under the theoretical framework of mobility biographies using non-traditional data, i.e., household panel surveys. Considering the availability of such data in multiple countries around the world, it is likely that more transportation-studies emanating from household panel surveys would appear and further proliferate in the coming years. One aspect that remains somewhat overlooked is the viability of using household panel surveys in the developing world to conduct transportation-related research. This dissertation aims to address this apparent gap using the case of Indonesia. To this end, the subsequent chapter elaborates on the research setting and overall research framework as the foundation for the three analytical chapters.

CHAPTER 3. RESEARCH CONTEXT AND FRAMEWORK

This chapter elaborates on the research context and the overall research framework that guides the analytical chapters of this dissertation. As the subsequent sections describe, a narrative of the research context is presented to substantiate the case as to why urbanizing Indonesia is the appropriate case to study the dynamic aspects of travel behavior. The theoretical framework, therefore, connects Indonesia's urbanization phenomenon with the theoretical proposition of mobility biographies to address the three research questions using a single primary longitudinal data source, i.e., the Indonesian Family Life Survey (IFLS).

3.1 Research Context: Urbanizing Indonesia

As briefly touched upon in the Introduction chapter, Indonesia offers a fitting case to examine travel behavior dynamics due to the evolving built environment and socioeconomic associated with rapid urbanization. This notion points to the assumption that the dynamics of travel behavior could be more observable and pronounced in Indonesia than, say, in places where urbanization was not as rapid as the one observed in the country. Examining travel behavior from the lens of rapid urbanization, therefore, warrant further analyses on Indonesia's urbanization process and characteristics. In doing so, an observation concerning comparative urbanization between Indonesia and fellow Southeast Asian countries, as well as with China and India, is presented in the following section. Subsequently, a within-country analysis of Indonesia's urbanization will be briefly discussed.

3.1.1 Comparative Urbanization

In this section, the analysis of comparative urbanization focuses on examining the urbanization rate, which is defined as the percentage share of a given country's population living in urban areas. In presenting the analysis, the emphasis is given on comparing Indonesia with fellow Southeast Asian countries (excluding Singapore due to its city-state status where the entire country is urbanized and Malaysia that has urbanization rate well above the rest of countries in the region), as well as China and India. The decision to include China and India in examining Indonesia's urbanization stems from the similar traits between these three countries in terms of population size, development status, and the Asian region these three countries are situated.

Figure 3 presents the urbanization rate of Indonesia, fellow Southeast Asian countries, as well as China and India from 1997 to 2017 using data from the World Bank – World Development Indicators (WDI) accessed through the 'WDI' package in R Studio platform (Arel-Bundock, 2019). As indicated, Indonesia has had the fastest urbanization growth in comparison to other Southeast Asian countries. During the last two decades, the share of Indonesians living in the country's urban regions rose from approximately 38.41% in 1997 to 54.66% in 2017. Other countries in the region that have had similar urbanization growth as Indonesia are Thailand and Vietnam. The data suggest that Thailand's share of urban population increased from 30.62% to 49.20% while Vietnam had seen its urban population rose from 22.96% to 35.21%.

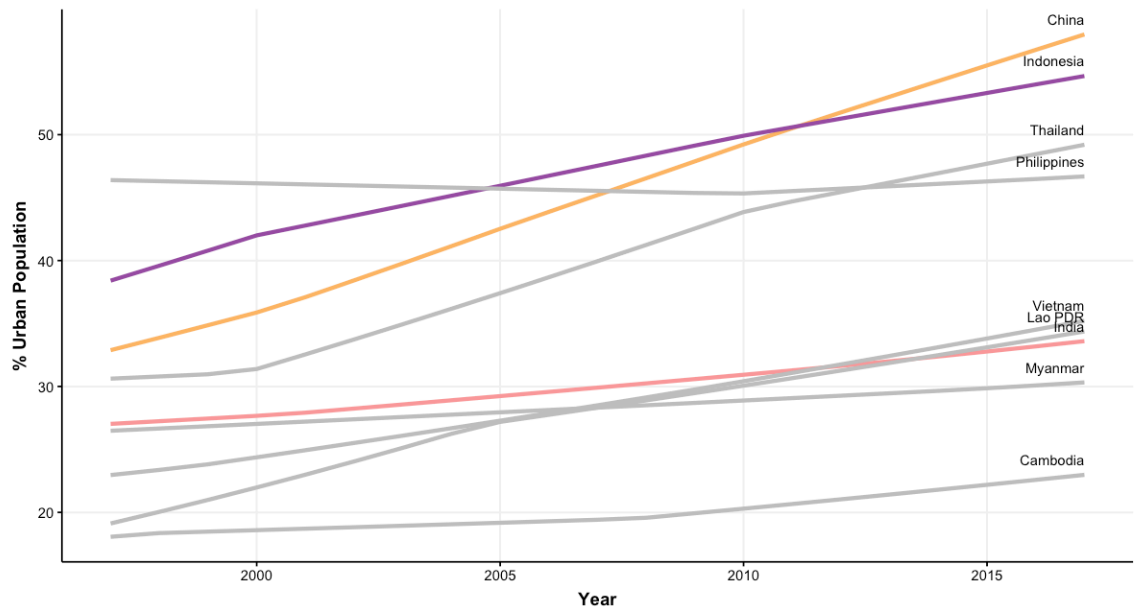

Figure 3 – Urbanization rate of Southeast Asian countries (excluding Malaysia & Singapore), China, and India

These figures of urbanization growth of Southeast Asian countries are, however, somewhat pale in comparison to the one observed in China. As indicated from the data, the share of the Chinese population living in the country's urban areas increased sharply from 32.88% in 1997 to 57.96% in 2017, or a 25.08% increase. During the same period, the growth of the urban population for Indonesia, Thailand, and Vietnam was 16.25%, 18.58%, and 12.25%, respectively.

In addition, the analysis also looks into the urbanization growth of India. An observation from the data and as indicated in Figure 3 suggests that India’s share of urban population rose 6.57% from 27.03% in 1997 to 33.60% in 2017, which is a noticeably slight increase in comparison to, for instances, China, Indonesia, and Thailand.

Figure 4 illustrates the extent of the urbanization rate in 2017. As shown, China and Indonesia are the only countries where more than half of their population residing in an urban environment. Moreover, taking into account the dashed vertical line that represents the average urbanization rate of the countries included in the analyses (40.55%), it is apparent that there were only four countries with above-average urbanization rate in 2017, i.e., China, Indonesia, Thailand, and the Philippines.

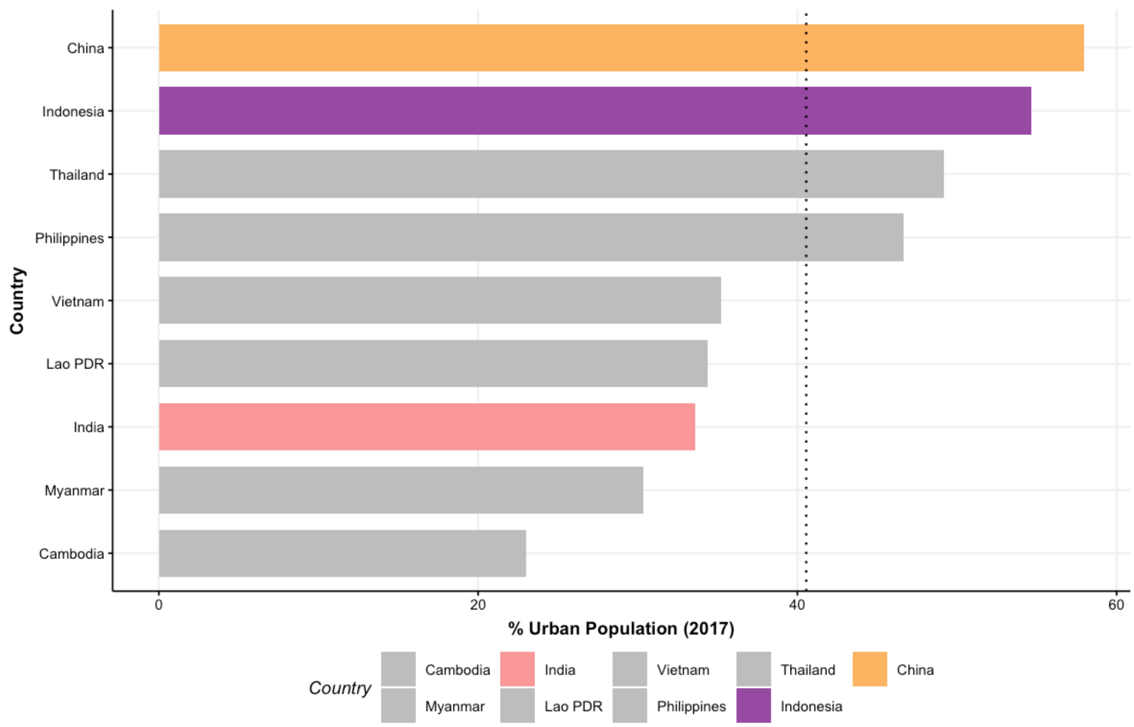

Figure 4 – Percentage of urban population (2017) of Southeast Asian countries (excluding Singapore and Malaysia), China, and India

The above analyses lend a hand to support the notion of rapidly urbanizing Indonesia in comparison to the neighboring countries and, therefore, the relative appropriateness in choosing Indonesia to examine travel behavior dynamics under the clout of rapid urbanization.

3.1.2 Three Driving Factors of Urbanization

A review of the urbanization process sheds light on the extent of the urbanization process in Southeast Asia and its neighboring countries, namely China and India. This section further elaborates on the three aspects that represent the urbanization process. That is, the urbanization process in this dissertation comprised of three driving forces: 1) rural to urban migration. 2) natural growth, and 3) reclassification (World Bank, 2018a).

Framing urbanization from these three factors, Figure 5 shows the extent of each factor in driving the urbanization process in Indonesia from 2000 to 2010, which is adopted from the World Bank's "Indonesia Economic Quarterly: Urbanization for All" report (Wai-Poi et al., 2018; World Bank, 2018a).

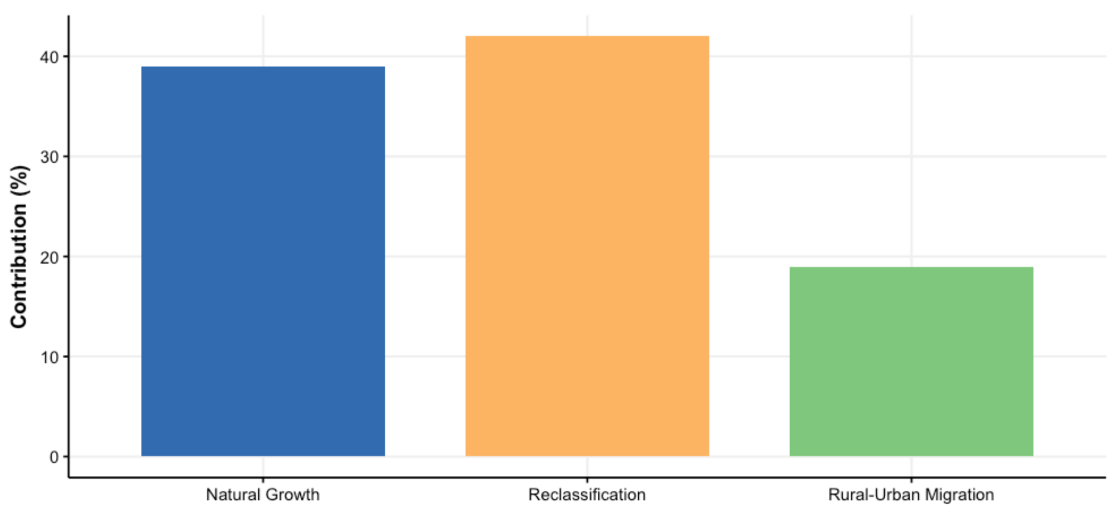

Figure 5 – The three driving factors of urbanization in Indonesia

(Source: World Bank (2018a), modified by author)

As indicated in Figure 5, the dominant factor influencing Indonesia's urbanization story is area reclassification. This aspect represents the process where a previously rural region turned into an urban due to a variety of factors, including densification, which made up approximately 42% of the country's urban transformation between the years 2000 and 2010. Natural growth, or densification of the country's existing urban fabric, contributed to around 39%. Rural to urban migration accounts for approximately 19% of the urbanization process in the country during that period.

This dissertation seeks to focus on rural to urban migration and natural growth to examine their association with travel behavior of the sample Indonesian households and individuals derived from the IFLS data. The subsequent section, therefore, further elaborates on the urbanization process in Indonesia, the overall research framework guided by the theoretical proposition of mobility biographies, and extended description of the IFLS as the primary data.

3.1.3 Urbanization in Indonesia

This section focuses on highlighting the uneven urbanization stage between regions in urbanizing Indonesia. A recent study posits that while Indonesia is rapidly urbanizing, the process is occurring unevenly where the Java region remains to be the prime urbanization hotspot in the country (World Bank, 2018a). To this end, these following sub-sections further highlight the extent to which Java dominates the urbanization process relative to other regions from the perspective of the share of urban population and employment by sector. Moreover, the subsequent section also discusses the interrelated dynamics between Indonesia's urbanization process and the country's evolving transportation landscape.

Population. The population distribution by region derived from the World Bank INDO-DAPOER (Indonesia Database for Policy and Economic Research) data, as shown in Figure 6, clearly illustrate the concentration of Indonesia’s population in Java. Figure 7 provides additional visual narratives of uneven urbanization. This figure shows population density at the district level in 2010 derived from INDO-DAPOER data based on the decadal 2010 Population Census. As illustrated, these two figures clearly show that the country’s population is disproportionately concentrated in Java. Indeed, every district in Java has population density well above the 3rd quartile (791.07 people/km^2 and 778.522 people/km^2 in 2010 and 2005, respectively) of all Indonesian districts.

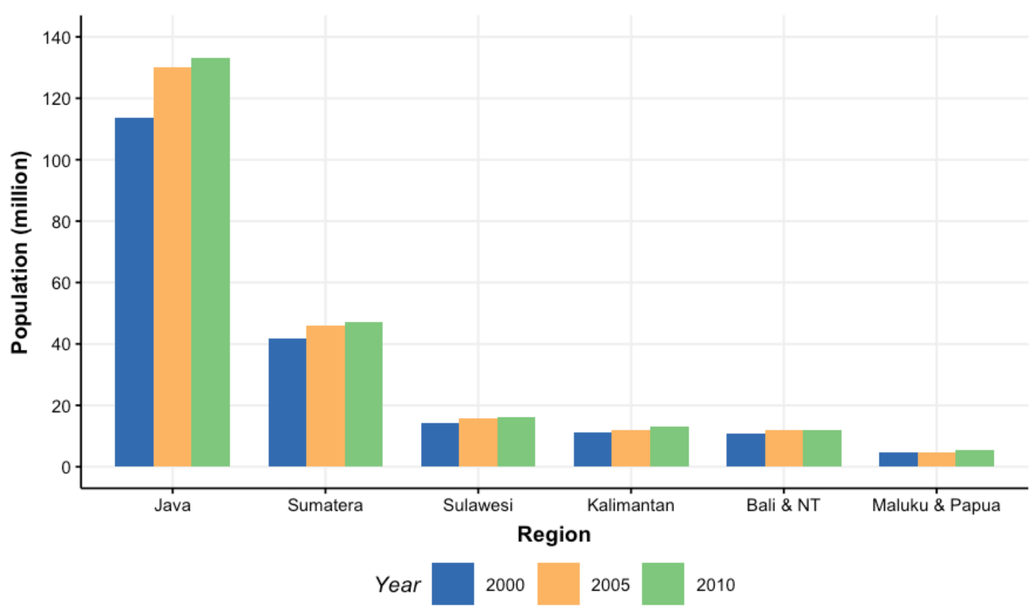

Figure 6 – Population distribution by region, 2000-2010

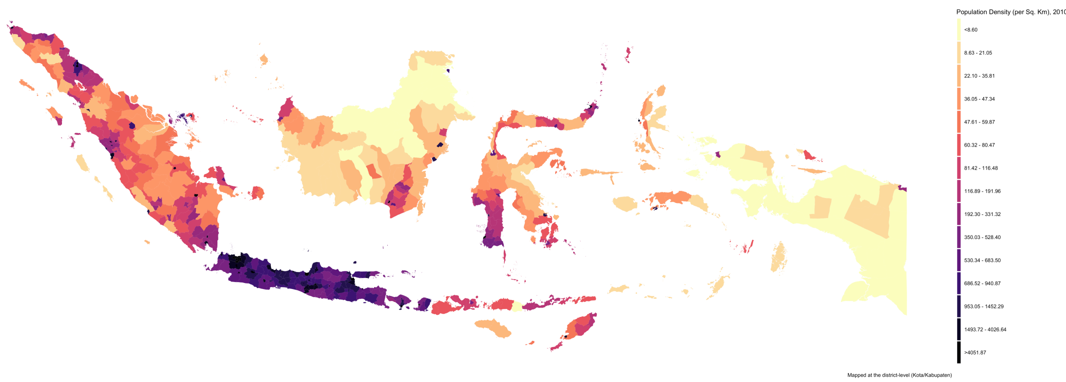

Figure 7 – Population density at the district level, 2010

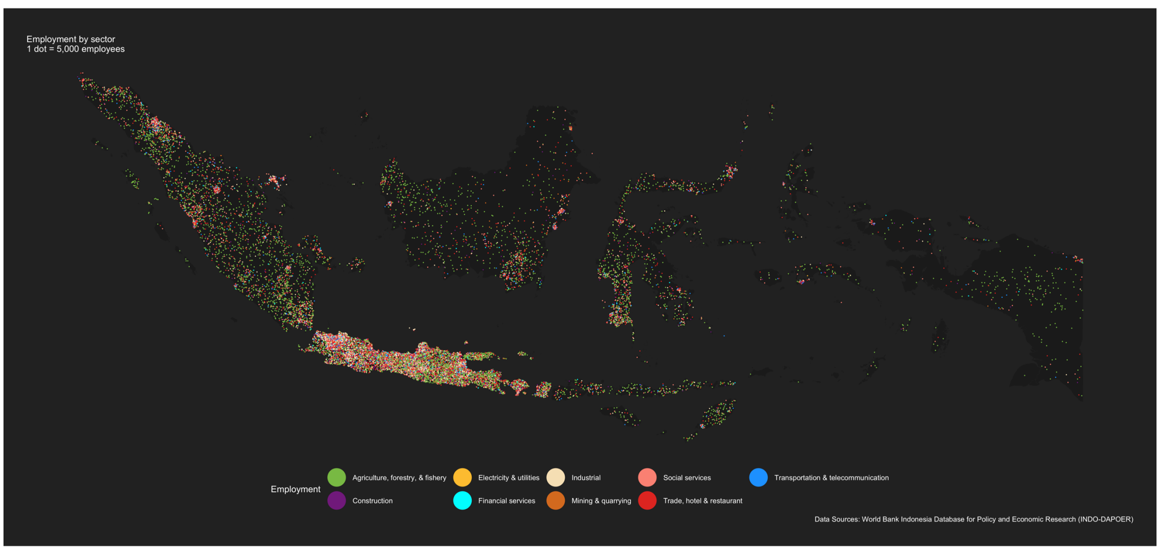

Figure 8 – Dot density map illustrating the distribution of employment by sector in Indonesia, 2015

Employment. A series of data visualization exercises are conducted using World Bank INDO-DAPOER data to shed further light on how employment is concentrated in Java. As shown in Figure 8, it is apparent how dense employment opportunities are in Java in comparison to, for instance, provinces in the eastern part of the country (e.g., Maluku, North Maluku, and Papua). Additional visualizations of employment by sector are presented in the Appendix.

Urbanization process and the evolving transportation landscape. Figure 9 and Figure 10 attempt to capture Indonesia's urbanization and development process and the country's transportation landscape over the past five decades using data derived from the World Development Indicators (Arel-Bundock, 2019; World Bank, 2018b) and Statistics Indonesia (Badan Pusat Statistik, 2017).

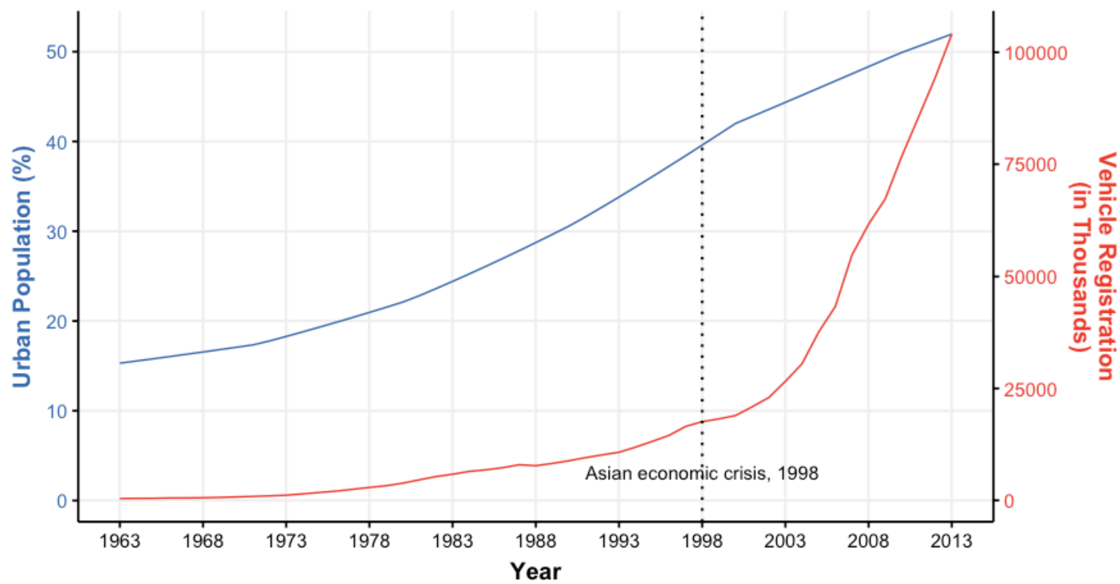

Figure 9 – Share of Indonesian population living in the country's urban areas overlapped with vehicle registration data, 1963-2013

Figure 9 shows how Indonesia is urbanizing as the share of the Indonesian population living in the country's urbanized areas grow over the past decades. This urbanization process appears to walk in tandem with vehicle registration growth, at least until late the 1990s. Starting from the early 2000s, however, the share of the urban population does little to explain the exponential vehicle registration growth in the country. While vehicle registration grew exponentially after the 1998 Asian economic crisis, the growth of the

Indonesian urban population had somewhat slowed down in comparison to the trend observed during the pre-economic crisis.

Figure 10 seems to paint a clearer picture that explains the country's evolving transportation landscape, especially during the period after the 1998 Asian economic crisis. That is, it appears that as the country recovered from the crisis, the trend of vehicle ownership per capita somewhat mimics the trajectory of GDP per capita.

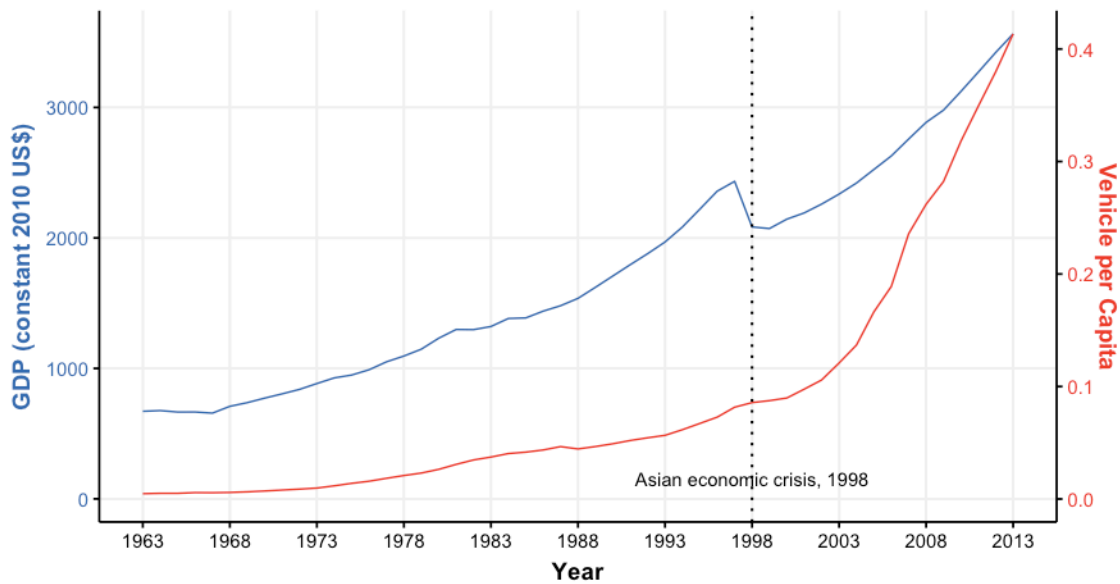

Figure 10 – Indonesia's GDP per capita (constant 2010 US$) overlapped with vehicle registration per capita, 1963-2013

3.2 Research Framework

Several factors that characterize the urbanization phenomenon drive the substantive topics this dissertation seeks to address, as illustrated in the research framework in Figure 11. As can be seen, embarking from the background phenomenon of urbanization, this dissertation will largely focus on two aspects that contribute to the urbanization process:

rural-urban migration and natural growth. This framework aligns with the theoretical proposition of mobility biographies that guide this dissertation, which asserts the relative importance of the dynamic aspect of the built environment through the emphasis on *accessibility domain* (Lanzendorf, 2003) and *residential biography* (Scheiner, 2007).

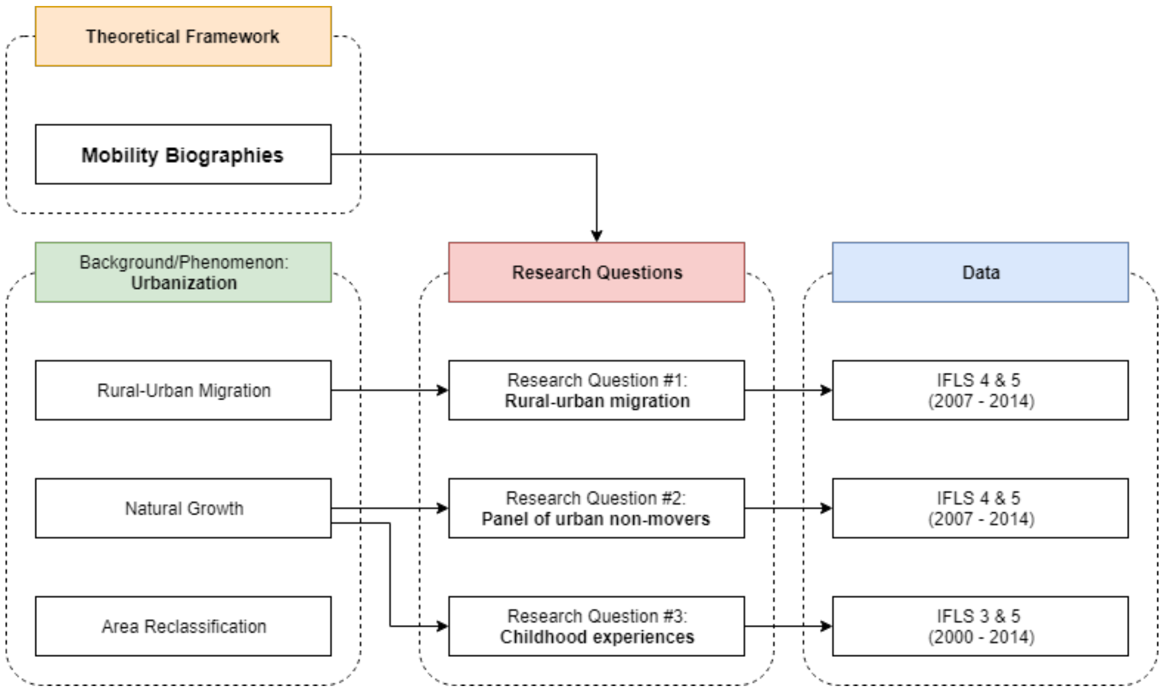

Figure 11 – Overall research framework

Following the research framework driven by the urbanization phenomenon and guided by the theoretical proposition of mobility biographies, as indicated in Figure 11, this dissertation focuses on addressing the three research questions. The first question revolves around the impact of relocating from rural to urban areas on travel behavior. The second question explores the evolving changes in the socioeconomic and built environment and its influences on travel behavior for urban non-movers. The third question aims to investigate the possible associations between childhood experiences on travel behavior, specifically walking habits during adulthood.

In sum, the overall research framework of this dissertation, as shown in Figure 11, specifies the linkage between the background phenomenon (i.e., urbanization process), the theoretical framework that guides the inquiries, the three research questions related to the background phenomenon, and the singular, primary source of data (i.e., IFLS - Indonesian Family Life Survey) to address the research questions. The following section will, therefore, further discuss the IFLS data, its characteristics, applicability for transportation-related research, as well as an overview of previous studies using this data.

3.2.1 Primary Data Source: Indonesian Family Life Survey (IFLS)

The Indonesian Family Life Survey (IFLS) is the longest, multitopic household panel survey in Indonesia, and arguably also among the longest outside of the OECD countries (Thomas et al., 2001, 2012; Witoelar, 2017). Initiated in 1993, the IFLS administrator has since then conducted four follow-up surveys in 1997, 2000, 2007, and 2014. Based on its baseline characteristics, IFLS is a nationally representative survey that represents approximately 83% of the country's population (Thomas et al., 2001, 2012; Witoelar, 2017).

One particular feature that somewhat highlights the notable characteristic of the IFLS is the relatively low attrition rate (Thomas et al., 2001, 2012). This characteristic is attributed to the survey design, where respondents who moved or relocated to different geographic locations are followed over time (Thomas et al., 2001, 2012). This survey design also helps explain the reasonably high recontact and survey completion rate that ranges from 90 to 95 percent (Witoelar, 2017).

3.2.1.1 <u>Viability of IFLS Data for Transportation Research</u>

As a multipurpose, general household survey, IFLS was not explicitly designed for transportation planning and research. Nonetheless, there are several indicators that can be used to conduct transportation-related research. This notion is evident in Figure 12 that depicts the similarities and differences between conventional transportation surveys and the IFLS. As can be seen, both forms of the survey provide similar socioeconomic indicators, while the geographic location of the surveyed households can be derived to construct built environment measures.

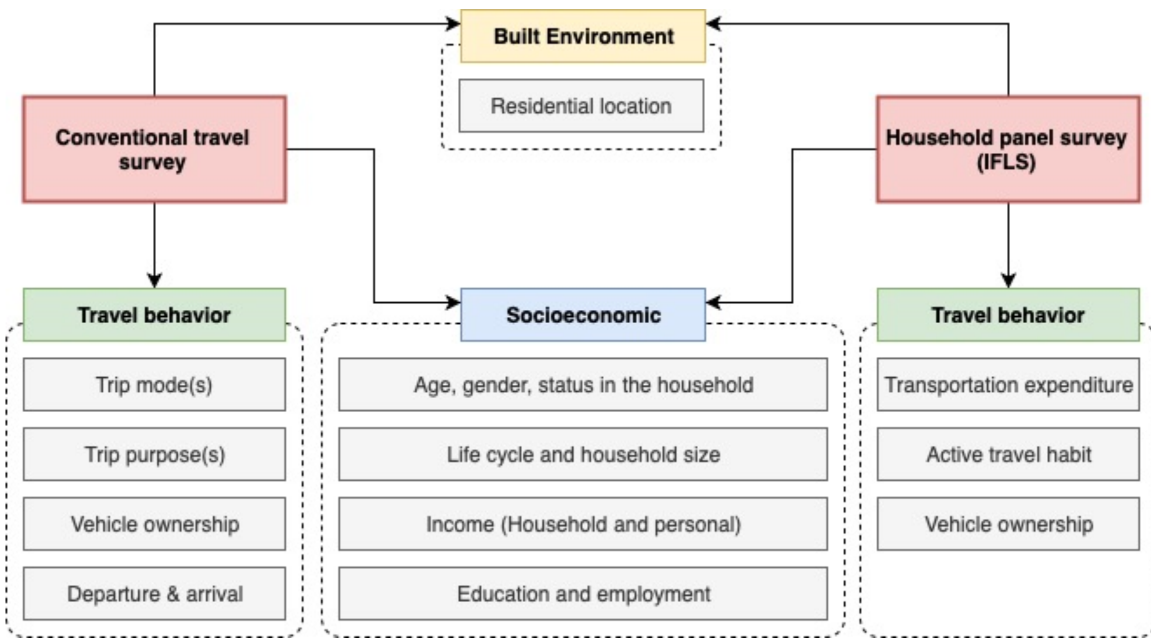

Figure 12 – Conceptual figure illustrating similarities and differences between conventional travel survey and household panel survey (IFLS)

The primary difference that separates these surveys revolves around the kind of travel behavior indicators each survey maintains. The IFLS, as is common in other household panel survey, does not offer detailed trips information that the typical travel survey would provide. Nonetheless, IFLS consistently probes information on household transportation expenditure (Table 5), which could be used as a proxy for more commonly-used travel behavior indicators, such as vehicle-miles traveled or trip frequency (Guerra, 2017; Guerra et al., 2018; Smart & Klein, 2018b). Table 5 also indicates the list of transportation-related indicators and their availability across survey waves.

Table 5 – List of transportation-related indicators and their availability across IFLS waves

Subject	**Question**	**Survey Wave**				
		IFLS 1	**IFLS 2**	**IFLS 3**	**IFLS 4**	**IFLS 5**
		1993	**1997**	**2000**	**2007**	**2015**
Non-Farm Business (NT) - Four-wheel motor vehicles	Does the household own [...] ?	✔	✔	✔	✔	✔
	What is the total (market) value of [...]?	✔	✔	✔	✔	✔
	What is the total value in rupiah of any [...] purchased in the past 12 months?	✔	✔	✔	✔	✔
Non-Farm Business (NT) - Other vehicles, specify:______*) i.e., 'two wheeler/motor cycle'	Does the household own [...] ?	✔	✔	✔	✔	✔
	What is the total (market) value of [...]?	✔	✔	✔	✔	✔
	What is the total value in rupiah of any [...] purchased in the past 12 months?	✔	✔	✔	✔	✔
Household Asset (HI/HR) - Vehicles	Do you or does any other member of the household own [...]?	✔	✔	✔	✔	✔
	What is the total value of [...]?	✔	✔	✔	✔	✔
	How many householders own [...]?	✔	✔	✔	✔	✔
	What is/was the tot. value in rupiah of any […] <u>purchased</u> in the past 12 mon?	-	-	-	✔	✔
	What is/was the tot. value in rupiah of any […] <u>sold</u> in the past 12 mon?	-	-	-	✔	✔
Consumption (KS) – Transp. expenditure	What were the total expenditures for […] during the past month…?	✔	✔	✔	✔	✔
Employment (TK) – Transportation benefits	Did you receive the following benefits from your employer for this job?					
	1. Car?	-	-	✔	✔	✔
	2. Transportation allowance?	-	-	✔	✔	✔
Health (KK) – Non-Motorized Transport	During the **last 7 days**, did you do any [….] for at least 10 min. continuously? / During the **last 7 days**, on how many days did you do [….]?					
• Moderate Phy. Activity	Yes: … No: … / …. Days	-	-	-	✔	✔
• Walk	Yes: … No: … / …. Days	-	-	-	✔	✔

Moreover, starting from IFLS 4, the survey administrator probed the questions surrounding a given individual walking behavior (Table 5). This indicator is a binary variable of whether the individual respondent walked for at least 10 minutes continuously but less than 2 hours in the past week. A discussion on the extent of how this dependent variable of interest is used in the literature is presented later on.

Prior to discussing the literature on the dependent variables of interest, a summary of studies that have used IFLS data is discussed in the following sub-section. The aim is to shed light on the applicability and viability of IFLS data to conduct planning-related research or studies that emphasize on the locational factors.

3.2.1.2 Relevant Studies using IFLS Data

This section highlights several studies using IFLS data that reasonably align with planning-related research. In selecting the reviewed studies, the emphasis is given on locational factors, which planners traditionally concern. For example, a somewhat recent study by Rosales-Rueda & Triyana (2018) explores the geographical variation of air pollution during the early stages of life and its effects on children's height during adulthood. They combined historical National Aeronautics and Space Administration's (NASA) Earth Probe Total Ozone Mapping Spectrometer (TOMS) for air pollution measurement and geographically linked it with the geocode information of IFLS respondents across waves.

Sujarwoto, Tampubolon, & Pierewan (2017) assess the varying factors influencing the quality of life and happiness using IFLS 4. While they did not apply a panel approach in their research design, the results of their study suggest that the locational factors appear to influence individual happiness. Specifically, they discover respondents who resided in districts with relatively good delivery and maintenance of public services reported happier and better quality of life than those who lived in districts with low-level of public service delivery (Sujarwoto et al., 2017).

Christiani, Byles, Tavener, & Dugdale (2015) use the fourth wave of IFLS to comparatively estimate the health characteristics of women living in major and smaller cities. While Christiani et al. (2015) did not utilize panel-based research design, their study reveals the relative influence of locational factors on women's health outcomes as measured in terms of blood pressure, body mass index (BMI), and tobacco consumption.

Gibson & Olivia (2010) exploit the longitudinal nature of the IFLS to estimate the effects of infrastructure on employment outcomes (non-farm employment, or NFE) of about 4,000 individuals in the survey. Specifically, they regressed the infrastructure characteristics (i.e., road and electricity) in two different years, 1993 and 1997, to estimate their influences on the likelihood of a given rural household to participate in NFE.

Several examples of studies that use IFLS as the primary data source, as described above, lends a hand to substantiate the relative high-quality of the IFLS data. This dissertation seeks to leverage the high-quality data the IFLS provides to address research questions that revolve around two primary dependent variables: household transportation expenditure (module KS, IFLS 4 and 5) and walking behavior (module KK, IFLS 5). The following section touches upon the present studies that have also used these two outcomes of interest.

3.2.1.3 The Literature on the Outcomes of Interest: Household Transportation Expenditure and Walking Behavior

Household transportation expenditure. The literature on the factors associated with household transportation expenditure is considerably small but growing. In each of the studies reviewed, the theoretical framework follows the proposition that expenses for transportation are primarily a function of socioeconomic and built environment factors. This framework is apparent in a 2017 study that examines transit expenditure using data from Mexico City (Guerra, 2017). In this study, Guerra (2017) specifies the socioeconomic factors as follows: income, household size, number of employed adults, and the average age in a given household. The expectation is that wealthier households would spend less on transit than an otherwise similar household but with lower income. Along a similar line, a higher number of children likely induces auto travel demand for most trips, thereby reduce transit expenditure. However, typical to most travel behavior and built environment studies, the inclusion of these socioeconomic indicators generally serve as control variables since the author focuses on estimating the relative influences of urban form factors, which include distance to the city center, population and employment density, and land use diversity. Consistent with the established literature on travel behavior and built environment results from a series of linear regression models suggest the inelastic relationship between built environment indicators and transit expenditure.

Along with similar research design and theoretical proposition, Guerra et al. (2018) assess the factors associated with household transportation expenditure in Buenos Aires using a cross-sectional dataset from the 2010 household travel survey in the region. Controlling for socioeconomic measures (e.g., household size, presence of the child, presence of adult workers, income, employment status, gender, and age), they were primarily interested to quantitatively estimate the influence of locational or urban form factors, which include land use diversity, population density, transit and car accessibility, on household transportation expenditure. Results indicate that socioeconomic factors exert a more pronounced influence than urban form factors. The modest impact of urban form factors is tractable from the indicator of population density. Specifically, for every 1% increase in population density, household transportation expenditure likely decreases by 0.2%, all else equal. Overall, results are consistent with the general findings in the literature on the relationship between urban form and distance traveled (Ewing & Cervero, 2010; Stevens, 2017), which is the most common outcome of interest.

Smart and Klein (2018a, p. 393) test the hypothesis that exposure to "more compact, transit-accessible, and walkable neighborhoods" would reduce transportation expenses. Using data from the PSID, the authors develop several regression models to test whether relocating to such an urban environment might reduce household transportation expenditure. Results indicate no meaningful impact of the relationship between dense neighborhoods served by reasonably ample transit services and transportation expenses. Instead, they argue that household transportation expenditure is by and large a function of socioeconomic traits. Findings suggest that as income, the number of adult workers in the household, and the number of children increase, transportation expenses might increase considerably as well (Smart & Klein, 2018a).

Walking behavior. In this dissertation, the indicator of walking behavior is derived from the binary answer to the query: "During the last 7 days, did you do any [*walk*] for at least 10 min. continuously?" (module KK, IFLS 5). The following section discusses to what extent does this type of indicator is used in the literature.

A 2017 study examines the relationship between walking (Wasfi et al., 2017). Based on the indicator of walking behavior derived from answers to the question "have you walked for exercise?", the authors estimated mixed-effects logistic regression models due to the nature of the dependent variable (Wasfi et al., 2017, p. 3). Following the theoretical proposition that travel behavior, including walking habit, is largely a function of the socioeconomic and built environment, the predictors included to model walking behavior reflect this theoretical framework.

Along a similar line, a host of other studies has modeled the factors associated with walking where the primary outcome is a self-reported binary indicator. For instance, a 2014 study exploits the variation of walking as a binary outcome derived from a self-reported survey in Australia (Villanueva et al., 2014). Berke et al. (2007) study the built environment correlates with a self-reported, binary indicator of the walking activity of older individuals in King County, Washington. A 2005 study also uses a self-reported indicator from a five-level Likert scale representing walking behavior from a survey in Oregon (Li et al., 2005).

In sum, the present studies lend a hand to support the use of the walking habit indicator derived from a self-reported survey to conduct relevant research. The walking behavior variable derived from the IFLS data aligns well with that proposition.

CHAPTER 4. TRAVEL BEHAVIOR EFFECTS OF RURAL TO URBAN MIGRATION

4.1 Introduction

Several factors drive the urbanization phenomenon, including the transition of places from rural to urban, densification as a result of natural population growth, and rural to urban migration, as indicated in Figure 5 (World Bank, 2018a). In this chapter, the focus is on rural-urban migration and its association with travel behavior changes. In conducting the analysis, this chapter shows a research design that exploits the prominent feature of the IFLS, i.e., tracking internal migration of the surveyed households and individuals over time or across survey waves. This feature allows analysis of the effects of relocating to an urban environment on travel behavior from a casual estimation perspective, which revolves around a before-after treatment-control evaluation design. In that vein, relocating households from rural to urban are assigned to a treatment group. In contrast, a subset of households who remained or relocated to other rural areas and have similarities with the treatment households, identified through Propensity Score Matching (PSM), is therefore assigned into a control group.

In the next section of this chapter, an overview of a small but growing literature on the travel behavior effects of residential relocation is discussed. Subsequently, a further elaboration of the data and methodologies is presented. The results and discussion section highlight the findings and its potential planning and policy implications as well as the contributions of this particular chapter to the travel behavior literature.

4.2 Literature Review: Residential Relocation

A survey of travel behavior literature focusing on residential relocation indicates that this strand of literature is considerably small but growing. A burgeoning interest on this subject stems primarily from a 2004 study that posits residential relocation as among the most influential determinant of changes in travel behavior (Klöckner, 2004). Using a stated preference survey, the author finds that 60.7% of the respondents surveyed consider relocating to a new city would likely change their daily mode choice (Klöckner, 2004).

One of the earlier explorations of travel behavior effects of residential relocation is a 2003 study in Seattle, WA (Krizek, 2003). Leveraging the longitudinal characteristic of Puget Sound Transportation Panel (PSTP), the study observes travel behavior changes of 430 households who relocated to a new residential location within the Puget Sound region. The relocating households are grouped with the entire sample, and regression models are developed to estimate the effects of residential relocation on VMT, PMT (Person miles traveled), number of tours, and number of trips per tour. The results indicate that greater neighborhood accessibility could lower VMT, PMT, and the number of trips per tour.

Considering the limited availability of longitudinal travel surveys, researchers have sought to estimate travel behavior effects of residential relocation using general-purpose household panel surveys. This is apparent in a 2007 study that leverages the longitudinal data of German Socio-Economic Panel (GSOEP) to analyze the connections among travel behavior, residential relocation, and life-course events (Prillwitz et al., 2007). A series of linear regression models are developed to estimate the factors associated with the changes in commute distance between 1998 and 2003. The findings indicate that several lifecourses events, e.g., job changes and residential relocation (i.e., relocated to the periphery), are statistically significant predictors of changes in commute distance.

A two-wave, before-after survey was developed to study the travel behavior effects of residential relocation in Beijing (Wang & Lin, 2017). The researchers visited several places, e.g., “real estate exchange centers, large furniture markets and home depots” (p.8), to identify households that were likely relocating to new residential locations. Of the approximately 500 households approached, 467 households were willing to participate and completed the first wave of the survey through a face-to-face interview. Of these respondents, 229 households (49.0% retention rate) completed the second survey wave. Results from cross-lagged panel regressions indicate that the built environment had modest influences on an individual’s travel behavior. Using the same data, a 2018 study focuses more on the roles of the social environment and personal networks in shaping travel behavior (Lin et al., 2018).

Under the circumstances when “true” panel data is not available, most studies examining the impacts of residential relocation on travel behavior typically employ retrospective interviews (Buchanan & Barnett, 2006; Fatmi & Habib, 2017; Klinger & Lanzendorf, 2016; Scheiner & Holz-Rau, 2013b; Stanbridge & Lyons, 2006). Travel behavior changes of newly relocated residents in the residential development of Northwood is the focus of a study conducted in Christchurch, New Zealand (Buchanan & Barnett, 2006). The authors investigate the present and past trip-making patters of 113 Northwood residents before and after moving to peripheral development through a retrospective survey. Their findings indicate that suburban-style development, such as Northwood, did little to alleviate widespread car use.

To tease out the travel behavior effects of residential relocation within the Beijing region, Yang uses data from the 1996 Household Relocation Survey (Yang, 2006). This study is among the few that employ the treatment-control approach. In this study, the author assigned respondents that relocated within the same districts as stayers or control group. In contrast, those who relocated across the districts within the city as treatment. The author recognizes that this approach might not be the ideal option since the control group had also relocated. Yang suggests that the ideal approaches would be "comparing pre-move and post-move travel behavior between the relocated households and comparing travel behavior between the relocated households and those who did not move" (2006, p. 10).

Unlike previous studies as described above that estimate travel behavior effects of residential relocation within a given city or region, the impact of long-distance, intra-city relocation across German cities, which include Hamburg, Bremen, Bochum, Dortmund, and Essen, was the centerpiece of a 2016 study (Klinger & Lanzendorf, 2016). The authors frame their analyses under the umbrella of analyzing travel behavior effects of exposure to different "mobility cultures" of each city. The primary argument of the study is that present literature on the relationships between residential relocation and travel behavior tends to overlook the worldwide phenomenon of long-distance moving since existing studies tend to focus simply on residential relocation within the region.

In sum, the literature on travel behavior effects of residential relocation is considerably small but gradually growing in recent years. Based on the extent of the literature review, there are several conceptual and methodological refinements that can be pursued and, therefore, could make contributions to the literature. First, assessing residential relocation effects on travel behavior would ideally be conducted through a longitudinal survey. Second, the longitudinal data would ideally be comprised of respondents who relocated to other places, as a treatment group, and those who maintained the same residential location across survey waves, as a control group. Moreoever, third, leveraging the longitudinal nature and treatment-control group of the data, causal analyses can be explored. These assessments largely reflect Yang's (2006) suggestions. This proposed research aims to incorporate these potential refinements by leveraging the panel nature of IFLS and employs a before-after, treatment-control evaluation design.

4.3 Conceptual Framework and Hypotheses

4.3.1 Conceptual Framework

From both theoretical and empirical perspectives, it is widely acknowledged that travel behavior is largely a function of socioeconomic and built environment factors. However, despite the fact that extensive empirical research has shown the extent of how these factors influence travel behavior, this theoretical foundation has not yet been tested, to the extent of the literature review, under the umbrella of socioeconomic and built environment changes over time using treatment and control group. This research seeks to fill this gap based on the conceptual framework, as presented in Figure 13, which takes cues from the 2010 study assessing the impact of rural-urban migration on health outcomes (Lu, 2010).

Specifically, Figure 13 shows that assessing the extent of socioeconomic and built environment factors on travel behavior dynamics can be examined from the lens of residential relocation. In this research, the assessment refers to the analyses of pre- and post-migration, where the baseline sample of rural residents was later separated into two groups: 1) rural to urban migrants are assigned into a treatment group while 2) non-migrants, or those who remained rural, are assigned into a control group. Following this identification strategy, this study expands the conventional approach in the present literature that mainly focuses on non-movers from a cross-sectional perspective. By analyzing travel behavior as a function of residential relocation, the conceptual framework, as shown in Figure 13, also highlights the contribution of this study within the literature.

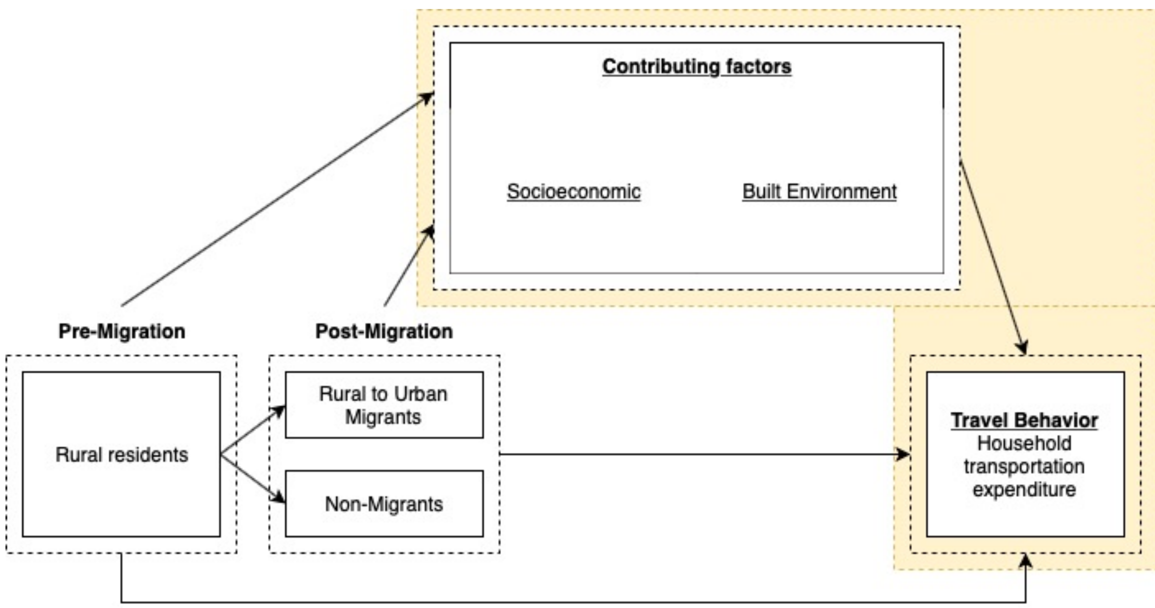

Figure 13 – Travel behavior as a function of residential relocation (rural to urban migration) and non-migrants

4.3.2 Hypotheses

Considering the present literature and the conceptual framework (Figure 13), the hypotheses of this research revolve around the expectation that the share of transportation expenditure from a total income would decrease as households relocated to urban areas. The urban environment understandably offers denser, more spatially diverse, and typically better infrastructure in comparison to rural areas. These attributes could cumulatively shorten travel distances, reduce travel costs, and decrease vehicle maintenance for a given urban household, therefore command lower share of transportation expenditure than an otherwise similar household living in a rural environment. Figure 14 supports the proposition that rural households spend more on transportation than households residing in urban areas.

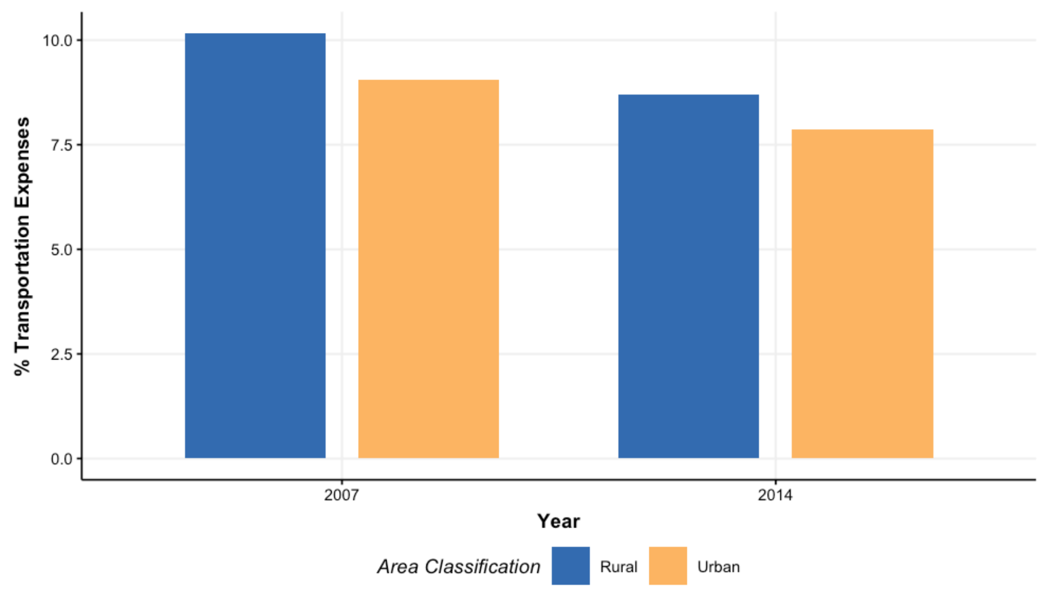

Figure 14 – Average percentage of transportation expenditure from a total income for IFLS 1) 2007 and 2) 2014

On the other hand, and notwithstanding the findings, as shown in Figure 14, relocating to urban areas could instead increase transportation expenditure for a relocating household. This proposition stems from the notion that the greater accessibility to a larger number of amenities in urban areas could induce a given household to make more trips and therefore spend more on transportation.

Given these likely scenarios as described above, the hypotheses and research question guiding this analytical chapter are as follow:

Q_1: How is a household's travel behavior affected by a move from rural to urban location?

H$_{\mathbf{Main}}$: Moving to urban areas reduces the share of transportation-related expenses as a result of greater density and a larger number of amenities by non-driving modes.

H$_{\mathbf{Alternative}}$: Moving to urban areas does not reduce the share of transportation-related expenses and could instead increase it since greater accessibility to a larger number of amenities could induce a given household to make more trips.

4.4 Data and Empirical Strategy

4.4.1 Data Structure

In addressing the research question and hypotheses, this study uses longitudinal IFLS data. The somewhat unique characteristic of the IFLS, where relocating households were tracked and interviewed, even if the household in question moved to different places, allows constructing a dataset that consists of treatment group, or respondents who relocated from rural to urban areas, and control group or respondents who remained in or relocated to other rural areas between IFLS 4 (2007) (Strauss et al., 2009) and IFLS 5 (2014) (Strauss et al., 2016).[2]

In identifying the sample for treatment group, the dataset identifies a subset of IFLS samples that satisfies these following conditions:

[2] As a note, IFLS adopt rural and urban classification from *Badan Pusat Statistik* (BPS) (Strauss et al., 2016). For a detailed description of urban-rural classification in Indonesia, see Badan Pusat Statistik (2010).

a) Moved out of the previous sub-district/*kecamatan*[3] (comparable to census tracts from the U.S. Census).

b) Conditional on *a*, the residential area classification of a given respondent has changed from rural (*desa*) in IFLS 4 into urban (*kelurahan*) in IFLS 5, and the population density at the sub-district/*kecamatan* level is greater than 1,000 people/km^2.

c) Information on household transportation expenditure as well as relevant socioeconomic indicators (e.g., household size, income) is available in both pre- and post-move.

Based on this identification strategy, two types of treatment group dataset are identified: 1) *split-off households* ($n_{\text{treatment}}$ = 263); and 2) *stem households* ($n_{\text{treatment}}$ = 30)[4].

Figure 15 presents the conceptual difference between stem and split-off households. As shown in the figure, the dataset for *split-off households* refers to new households captured in IFLS 5 that was formed by a portion of household members originated from the rural origin households as recorded in IFLS 4. In this study, a subset of households that resided in urban areas in IFLS 5 from the dataset for split-off households was assigned into the treatment group. The treatment designation for a subset of households in the dataset for *stem households* refers to origin households that moved to urban areas where none of its

[3] IFLS 5 (Strauss et al., 2016) categorize several types of residential relocation: 01 ___.___ kilometer in the same village (*desa/kelurahan*); 11 Move out of the village, same sub-district (*kecamatan*); 12 Move out of the village, same district (*kabupaten*); 13 Move out of the village, same province; 14 Move out of the village, different province. Due to the fact that geocode information is only available at the sub-district-level, this study excludes respondents who relocated within the same sub-district.

[4] The initial treatment group for both datasets has more observations than the working datasets. However, several indicators and the outcome variables (e.g., household composition, income, transportation expenditure) are found missing, thereby reduce the number of observations in each dataset. This is understandable considering the complexity and a wealth of information accumulated in the IFLS.

household members are found to form a new household as documented in the IFLS. For reference, in the IFLS household crosswalk data, the split-off households are coded as '1.splitoff' while stem households are coded as '0.stem'.

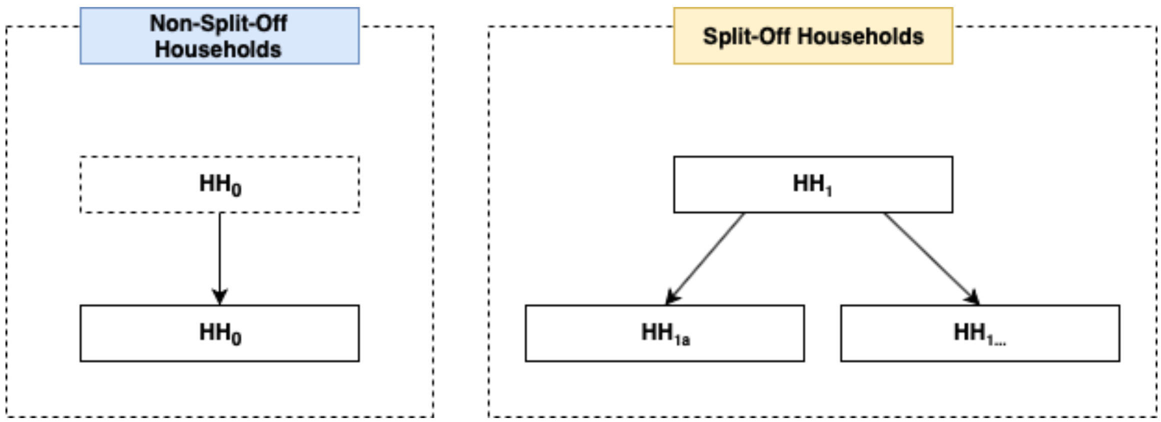

Figure 15 – Conceptual difference between stem and split-off households

While the treatment group is identified as discussed in the preceding section, the data does not automatically provide the control group. The notion that IFLS is not specifically and exclusively designed for this study is the contributing factor that explains this challenge. Toward this end, Propensity Score Matching (PSM) is applied to the original dataset as a means to construct the control group, or a group of households who were identified as having similar characteristics with the treatment group, as described in the following section.

4.4.2 Propensity Score Matching and Selection Model

4.4.2.1 <u>Propensity Score Matching</u>

The application of the PSM approach is relatively extensive, especially for evaluating social programs (Ravallion, 2008). The application of PSM is particularly relevant when working with observational data or when randomization was not possible, or under the condition where the identified control group(s) does not exhibit sufficient similarities with the treatment group, known as randomization failure (Austin, 2011; Bonell et al., 2011; Olmos & Govindasamy, 2015).

Application of PSM means developing the selection model to ascertain treatment and control group have similar traits. Several studies have discussed the theoretical and methodological perspectives on the choice of explanatory variables to be included in the selection model (Austin, 2011; Brookhart et al., 2006; Olmos & Govindasamy, 2015; Starks & Garrido, 2014). Among the studies reviewed, it appears that variables selection in the propensity score selection model should be primarily informed by the theory or literature rather than simply observing the p-values of the variables predicting membership in either treatment or control group using logistic regression (Brookhart et al., 2006; Starks & Garrido, 2014).

A multitude of studies indicates the growing popularity of PSM techniques in travel behavior research. Travel behavior researchers typically employ this technique as a means to compare outcomes between two similar groups. For instances, activity-travel pattern between residents of private and public housing in Hong Kong (Wang & Cao, 2017), vehicle use for shopping between recent-movers and existing residents in suburban Atlanta, GA (Y. Lee & Guhathakurta, 2018), VMT between residents of Transit-Oriented Development (TOD) neighborhoods and non-TODs (Nasri et al., 2018; Park et al., 2018), travel behavior between urban and suburban dwellers in Boston, MA (J. S. Lee et al., 2014), among others. From several studies as described, it has become apparent that the PSM technique in travel behavior research typically involves incorporating sociodemographic variables in the selection model. As indicated in a number of studies summarized earlier, these include household size, employed household members, income, vehicle ownership as the most common variables used in the studies reviewed (J. S. Lee et al., 2014; Nasri et al., 2018; Park et al., 2018).

4.4.2.2 Selection Model

Considering the literature and the data availability from the IFLS, the following variables are incorporated in the selection model for both stem and split-off households: household size, number of employed household members, number of children (household members <5 years old), income, and vehicle ownership. Before discussing the procedure to run the selection model and evaluate the treatment and control balance, a qualitative observation that shows the pre-matching difference between treatment and control group is presented in Figure 16 and Figure 17 for stem and split-off households, respectively.

Based on a qualitative observation from both figures, it might be inferred that relocating rural-urban households are typically younger, smaller in terms of household size, and had a slightly higher income than the control households. This observation is, to some extent, fits with the expectation. For instance, a given rural household that maintains a considerably large household size might be more unlikely to relocate to urban regions than similar households but with smaller household sizes due to the likely substantial resources required to do so. It is also understandable that relocating households tend to be younger than those who remained rural since younger households might be more willing to take the risks of moving to new places.

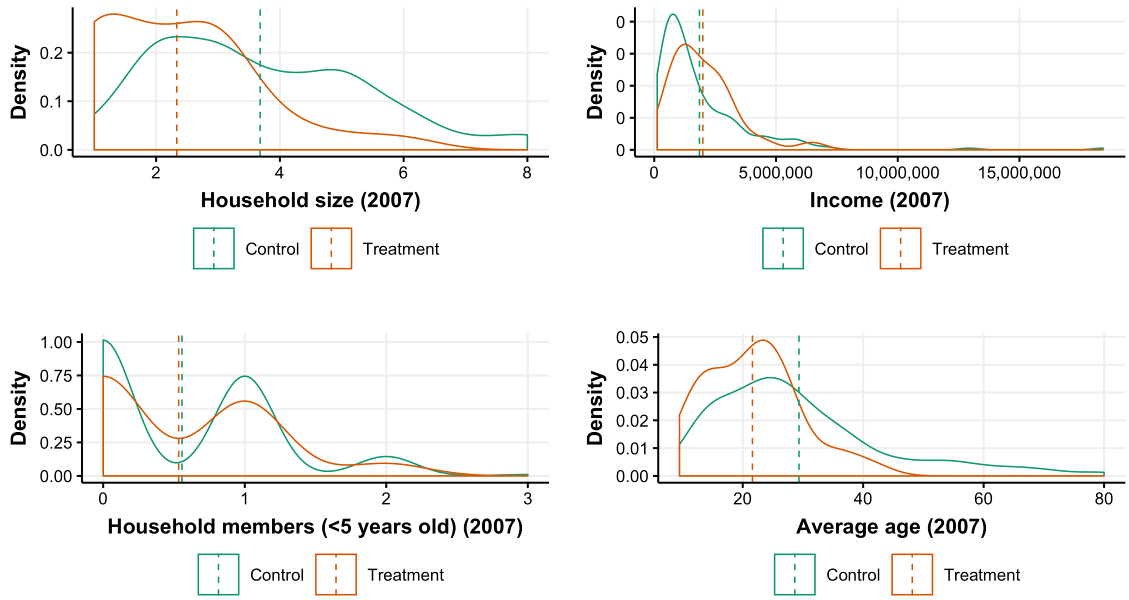

Figure 16 – Pre-matching comparison between treatment and control households, stem households

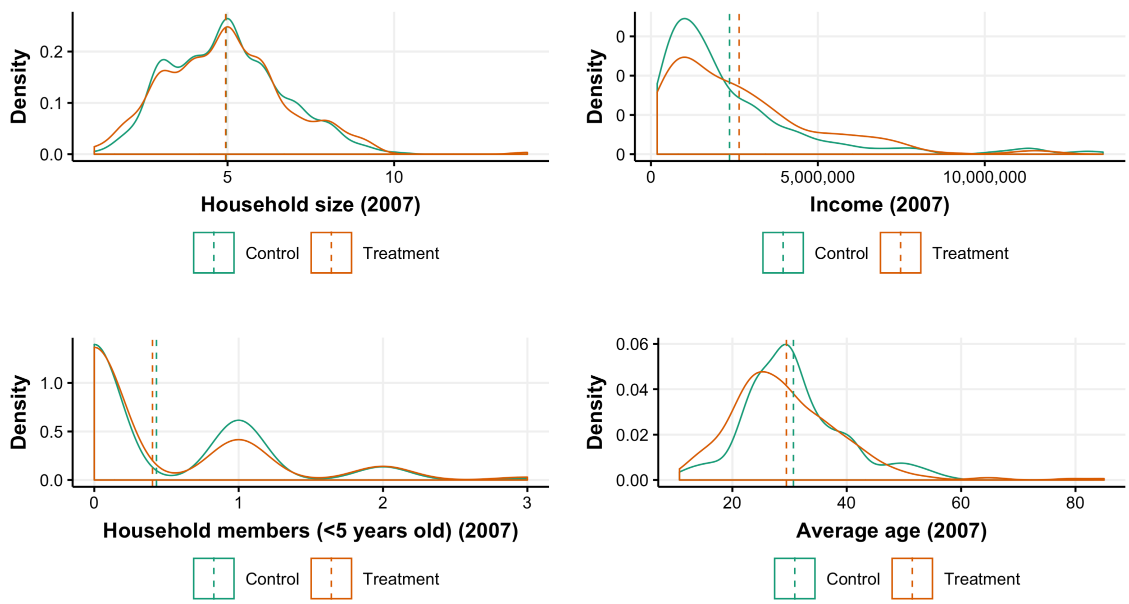

Figure 17 – Pre-matching comparison between treatment and control households, split-off households

Using the variables as listed above to run the matching process, the selection model is estimated using the *MatchIt* package (Ho et al., 2018) in the R Studio platform (R Core Team, 2013). Several matching procedures offered in the package are explored, including *nearest*, *optimal*, and *genetic*. The behavior of each matching procedure and the resultant matched samples are evaluated. The assessment of the resulting matched samples indicates that *optimal* matching would be the more appropriate option than *nearest* and *genetic* matching. In addition to evaluating the resultant matched samples, it should also be noted that *genetic* matching is also not recommended for a small dataset ($n < 1000$) (Diamond & Sekhon, 2006). Due to these reasons, in both the selection model for the stem and split-off households as presented below, the results refer to the ones produced from the *optimal* matching.

4.4.3 Evaluating Covariates Balance

The literature on PSM suggests that evaluating the covariates balance in the matched and unmatched dataset can be done through two primary methods: 1) numerical evaluation of the mean and standardized differences between treatment and control group (Austin, 2009); and 2) graphical assessment to ascertain the presence of *common support assumption*, or the notion that the treatment and control group overlap and have similarities from one to another (Caliendo & Kopeinig, 2008). The graphical assessment can be conducted using "a visual analysis of the density distribution of the propensity score in both groups" (p.12).

The numerical evaluation based on standardized differences follows Equation 1 below (Austin, 2009). As indicated, the standardized difference (d) of a given variable is the product of the mean of the treated subjects ($\bar{x}_{treatment}$), mean of the untreated subjects ($\bar{x}_{control}$), variance of the treated subjects ($var_{treatment}$), and variance of the untreated subjects ($var_{control}$) in the respective variable.

$$d = \frac{(\bar{x}_{treatment} - \bar{x}_{control})}{\sqrt{\frac{var^2{}_{treatment} + var^2{}_{control}}{2}}} \tag{1}$$

For the stem households, the initial dataset constructed from the IFLS consists of 30 households in the treatment group or households who relocated from rural in 2007 to urban in 2014, and 185 households in the control group, or a subset of households that lived in the same districts as the treatment households and remained rural. For the split-off households, the initial dataset comprised of 263 treated households group, or household members who relocated from their originating rural households and formed a new household in urban areas, and 302 households in the control group, or household members who relocated and formed a new household in other rural locations from their originating rural households. Using these two datasets, the PSM process, as specified in the preceding section, is applied by specifying a 1:1 ratio between treatment and control; thus, the resultant households in the control groups correspond with the number of treated households. That is, the stem dataset now consists of the same number of households in the treatment (n=30) and control group (n=30, from the initial 185 households). Similarly, the dataset for split-off households is now comprised of 263 households in the treatment group and 263 control households, from the initial 302 households in the control group.

To evaluate how results from the matching process improved the similarities between households in the treatment and control group, a comparison between unmatched and matched samples is shown in Table 6 and Table 7 for the stem and split-off households, respectively. For the stem households, as shown in Table 6, the matched dataset has a considerably better covariates balance between treatment and control group than the unmatched one. In almost every variable, the standardized difference falls below the 0.25 threshold, which ascertains that the matching process yields a more balanced dataset (Austin, 2009). However, it appears that one variable in the matched dataset, i.e., employed household members, has a standardized difference that is slightly larger than the suggested 0.25 threshold. Nonetheless, this difference is a considerable improvement from the unmatched dataset, where the standardized difference was 0.95.

Table 6 – Covariates balance: Before and after matching between treatment and control, stem households

Variables (1)	**Matched Unmatched** (2)	**Mean Treatment** (3)	**Mean Control** (4)	**Mean Difference** (5)	**Standardized Difference** (6)
Age	U	21.61	29.34	-7.72	-0.05
	M	21.61	20.12	1.50	0.03
Income	U	1.994.130,73	1.856.160,48	137.970,26	0.00
	M	1.994.130,73	1.895.281,03	98.849,70	0.00
Household size	U	2.33	3.68	-1.35	-0.56
	M	2.33	2.50	-0.17	-0.11
# Children <5 years old	U	0.53	0.56	-0.02	-0.06
	M	0.53	0.50	0.03	0.10
Employed HH members	U	0.87	1.58	-0.71	-0.95
	M	0.87	1.00	-0.13	0.30
Vehicle ownership	U	1.13	0.99	0.14	0.17
	M	1.13	1.10	0.03	0.04

Note: U = Unmatched; M: Matched

The results from the matching process illustrating the comparison between unmatched and matched samples for the split-off households are presented in Table 7. As indicated in the standardized mean difference column, the matched dataset has a better covariates balance than the unmatched dataset. This notion is particularly evident as the standardized difference of each variable in the matched dataset has a lower value than in the unmatched dataset. To further support the notion that the matched dataset yields similar traits between treatment and control samples, the standardized difference between treatment and control in each variable falls below the suggested 0.25 threshold (Austin, 2009).

Table 7 – Covariates balance: Before and after matching between treatment and control, split-off households

Variables (1)	**Matched Unmatched** (2)	**Mean Treatment** (3)	**Mean Control** (4)	**Mean Difference** (5)	**Standardized Difference** (6)
Age	U	29.45	30.67	-1.23	-0.01
	M	29.45	30.06	-0.61	-0.01
Income	U	2.641.639,33	2.353.640,24	287.999,09	0.00
	M	2.641.639,33	2.494.268,31	147.371,02	0.00
Household size	U	4.94	4.96	-0.01	0.00
	M	4.94	4.94	0.00	0.00
# Children <5 years old	U	0.40	0.43	-0.03	-0.06
	M	0.40	0.41	-0.01	-0.03
Employed HH members	U	2.06	2.21	-0.14	-0.11
	M	2.06	2.09	-0.02	0.02
Vehicle ownership	U	1.32	1.03	0.29	0.26
	M	1.32	1.14	0.17	0.16

Note: U = Unmatched; M: Matched

In addition to assessing the matching through numerical evaluation, the evaluation of the resultant covariates balance from the perspective of the region of common support between treatment and control group is shown in Figure 18. This approach revolves around visually inspecting to what extent treatment and control group overlaps. As indicated, the overlapped area between these two groups is visible and thus ascertain the substantial similarities between treatment and control households in both datasets for stem and split-off households.

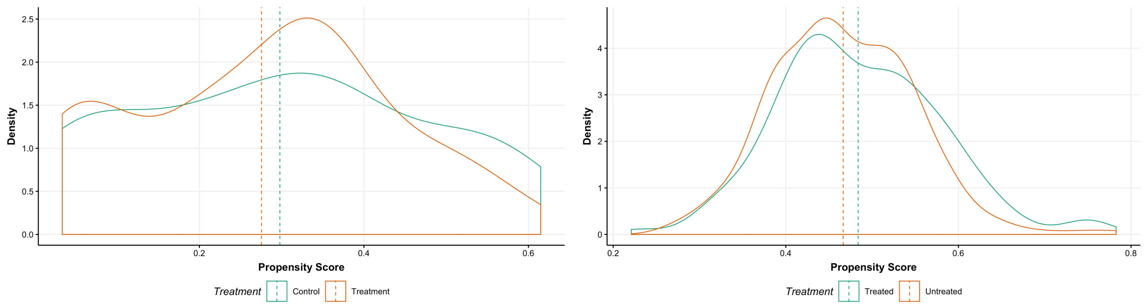

Figure 18 – Density function of propensity scores between treatment and control samples, stem households (left) and split-off households (right)

The improved covariate balance, as shown above, suggests that the matching would enhance the internal validity (Lim et al., 2014; Stuart et al., 2001). Along this line, scholars have also lamented how the matching technique interacts with the external validity, particularly as several respondents were excluded after the matching (Lim et al., 2014; Stuart et al., 2001). In regard to this, and considering the empirical strategy as described in the subsequent sub-section, this study follows the proposition as suggested by Kahn-Lang and Lang (2019) on the importance of baseline or pre-treatment similarity between treatment and control and thereby the appropriateness of the matching (Kahn-Lang & Lang, 2019; McKenzie, 2020).

4.4.4 Migration Pattern

Looking into relocating categories in both datasets for the stem and split-off households, as shown in Table 8, it indicates the apparent difference in terms of migration destination between the two groups. For the stem households, the majority of the samples (53.3 percent) relocated to urban areas in the same province – comparable to states in the

U.S, 40 percent relocated to urban areas within the same district (*kabupaten*). In comparison, the remaining 6.67 percent relocated to urban areas outside of the originating province. This pattern indicates that when the whole members of a given household relocated to a new residential location, they tend to choose locations that are somewhat nearby from the initial location.

A slightly different pattern is observed for the split-off households. As can be seen in Table 8, 17.87 percent of the samples in the dataset for split-off households relocated to urban areas in the same district, 47.52 percent relocated to urban areas in a different district, but the same province, while the remaining 34.61 percent relocated to a different province.

Table 8 – Characteristics of rural to urban migration

	Stem		Split-Off	
	Number	Percent. (%)	Number	Percent. (%)
Move out of the village, same district (*kabupaten*)	12	40.00	47	17.87
Move out of the village, same province	16	53.33	125	47.52
Move out of the village, different province.	2	6.67	91	34.61
Total	30	100	263	100

Taken together, the somewhat contrasting migration pattern between the stem and split-off households suggests that the decision to relocate an entire household (i.e., stem households) to distant urban regions, particularly the ones located in a different province, likely entail substantial efforts that most rural households could not afford. This notion partly explains why the ones who decided to relocate chose to do so in places close to the initial location.

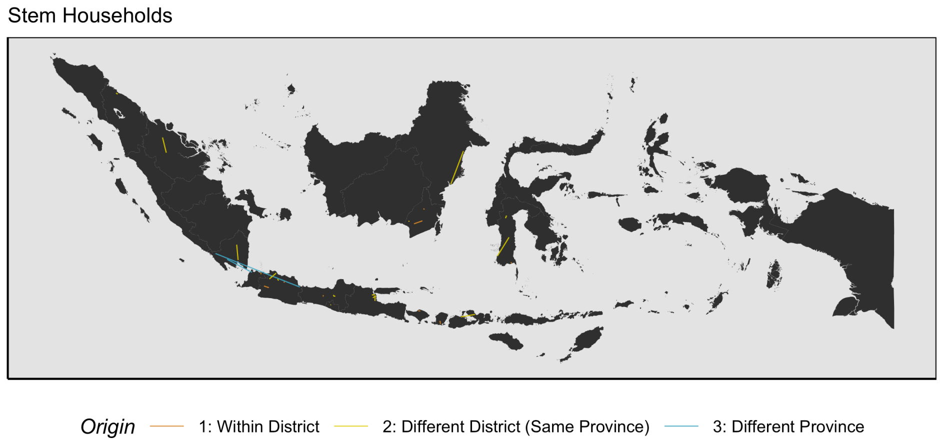

Figure 19 – Rural to urban migration, stem households

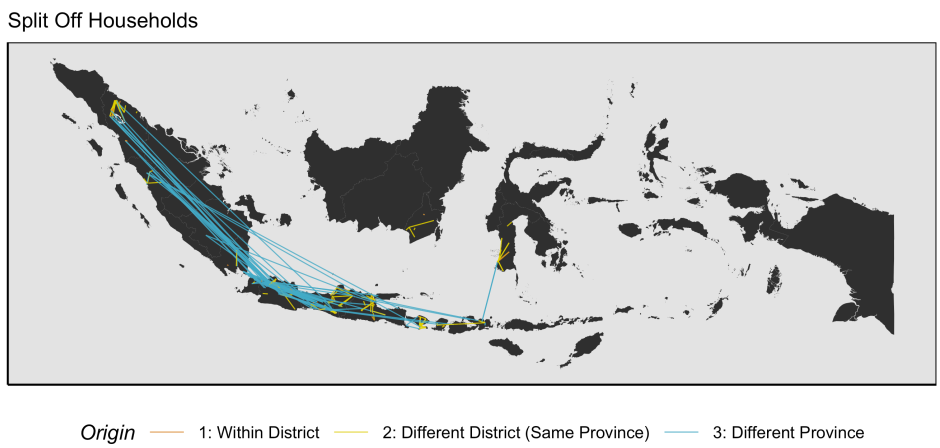

Figure 20 – Rural to urban migration, split-off households

For an illustrative purpose, Figure 19 and Figure 20 show the directional flows of rural to urban migration for the stem and split-off households, respectively. Following the migration classification from the IFLS 5 (Strauss et al., 2016), the orange line indicates 'Move out of the village, same district (*kabupaten*)', the yellow line indicates 'Move out of the village, same province', and the blue line indicates 'Move out of the village, different province'. Due to the considerably small samples of the dataset for stem households (Figure 19), the directional flow, as shown in the figure, might not be too apparent. The flow appears to be far more apparent in the figure for the split-off households (Figure 20). As can be seen, visual observation of the figure indicates that most relocating households moved to Java island, particularly to the Greater Jakarta and its surrounding region.

Table 9 reaffirms the visual observation of migration flow, as shown in Figure 20, by indicating the percentage distribution of destination for a subset of split-off households who relocated to different provinces. As indicated, it is evident that most rural to urban migrants relocated to Jakarta and its surrounding regions, i.e., Banten and West Java. Indeed, almost two-thirds of the samples (i.e., 70.33%) relocated to municipalities in these regions. This finding is expected considering the relative economic power and, therefore, the attractiveness of Jakarta and its surrounding provinces as a prime migration destination, at least as suggested by the samples from the IFLS data.

Table 9 – Provincial destination of rural-urban migrants, split-off households

Province	Region	2014		
		# Households	**%**	**% (by region)**
Bali	Bali & NT	3	3.30	6.60
West Nusa Tenggara	Bali & NT	3	3.30	
Banten	Java	11	12.09	85.70
Central Java	Java	1	1.10	
East Java	Java	6	6.59	
Jakarta (Capital)	Java	26	28.57	
West Java	Java	27	29.67	
Yogyakarta	Java	7	7.69	
Lampung	Sumatera	1	1.10	4.40
North Sumatera	Sumatera	2	2.20	
West Sumatera	Sumatera	1	1.10	
South Kalimantan	Kalimantan	1	1.10	1.10
South Sulawesi	Sulawesi	2	2.20	2.20
	Total	91	100	

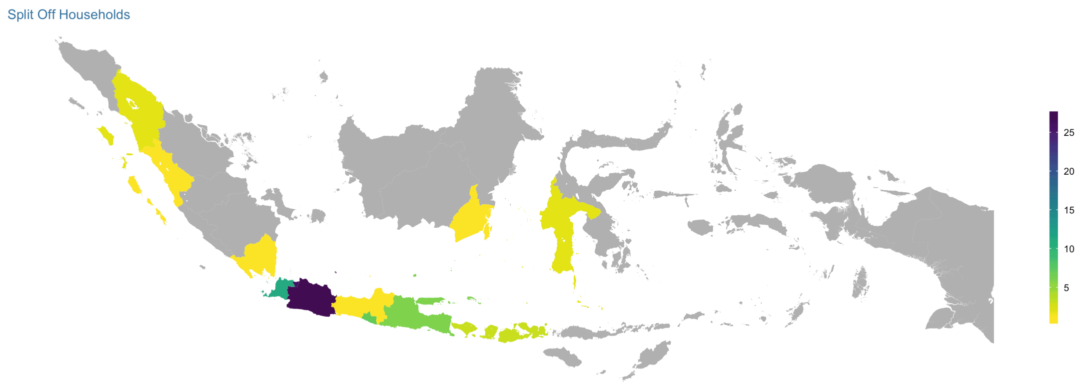

Figure 21 – Provincial destinations of rural-urban migrants, split-off households

4.4.5 Empirical Strategy

Following the overall research design that revolves around before-after, treatment-control approach, as indicated in the previous section, this section further elaborates the empirical strategy to estimate the effects of relocating to urban areas on transportation expenditure using Difference-in-Differences (DID) approach. As the name suggests, the DID estimator involves computing **before** and **after** *difference*. In this context, the **before** *difference* refers to the value of computing the average transportation expenditure ($\bar{y}$) between a subset of samples which eventually relocated ($\bar{y}_{T,B}$) to urban areas and a subset of samples who did not relocate and remain in rural areas ($\bar{y}_{C,B}$). This approach is represented in the following equation: $\bar{y}_{T,B} - \bar{y}_{C,B}$. As a note, considering the propensity score as shown in the previous section yields noticeably similar samples between the treatment and control households, concerns over selection bias that might bias the DID estimator are minimized.

Similarly, the **after** *difference* refers to the value of computing $\bar{y}_{T,A} - \bar{y}_{C,A}$, or the difference of average transportation expenditure ($\bar{y}$) between a subset of samples who had relocated to urban areas ($\bar{y}_{T,A}$) and those who remain in their respective rural areas ($\bar{y}_{C,A}$). Considering these two differences, the DID estimator ($\delta^{\wedge}$) is therefore computed as in this following Equation 2:

$$\delta^{\wedge} = \left(\bar{y}_{T,A} - \bar{y}_{C,A}\right) - \left(\bar{y}_{T,B} - \bar{y}_{C,B}\right) \tag{2}$$

The DID approach, as shown above, can be extended into a linear regression estimation. This way, a range of factors that might influence the outcome of interest can be included in the equation. At a minimum, however, the DID model would include variables indicating the *treatment* (T), a binary indicator where 1 represents treatment group and 0 represents non-treatment or control group; *time* (A), which is also a binary indicating *before* and *after*; and the interaction term between *treatment* and *time* as the primary variable of interest (δ) as shown in Equation 3. Other covariates can be incorporated accordingly.

$$y = \beta_0 + \beta_1 T + \beta_2 A + \delta_0 TxA + e \tag{3}$$

In this research, several other socioeconomic covariates (SE) are further incorporated as exposure to the dense and diverse built environment in urban areas in contrast to rural areas might not, by itself, exert significant influences on household transportation expenditure. The inclusion of these socioeconomic variables (SE) is shown in Equation 4.

$$y = \beta_0 + \beta_1 T + \beta_2 A + \delta TxA + \beta_3 SE + e \tag{4}$$

4.5 Results

4.5.1 *Changes in average transportation expenditure (t-test)*

Before discussing results from DID models, the comparison table illustrating the changes in the outcome of interest, i.e., the share of household transportation expenditure from a total income, between treatment and control group for both stem and split-off households is presented in Table 10. As can be seen, this descriptive statistic does seem to suggest that relocating households to urban areas had a greater reduction in transportation expenditure than a similar group of households that remained rural.

Table 10 – Comparison of average transportation expenditure

Dataset	Type	Period	Household Transportation Expenditure: Avg. Transp. Expenditure	Average Changes	t-test (*p-value*)
Stem Households	Control	2007 (Rural)	0.126	-0.004	0.908
		2014 (Rural)	0.123		
	Treatment	2007 (Rural)	0.131	-0.052	0.043**
		2014 (Urban)	0.079		
Split-Off Households	Control	2007 (Rural)	0.112	-0.008	0.364
		2014 (Rural)	0.103		
	Treatment	2007 (Rural)	0.130	-0.020	0.059*
		2014 (Urban)	0.109		

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Furthermore, this notion is apparent for both types of migrants, i.e., split-off and stem households. Specifically, for stem households, it appears that households that did not relocate virtually maintain the same transportation expenditure from a total income, i.e., 12.6 and 12.3 percent in 2007 and 2014, respectively (*p-value* = 0.91). On the other hand, households that relocated to urban regions appear to have their share of transportation expenditure from a total income reduced by, on average, 5.2 percent (*p-value* < 0.05).

A somewhat similar trend is apparent in the dataset for split-off households. It appears that relocating households have their transportation expenditure from a total income reduced by approximately 2 percent (*p-value* < 0.1), while the reduction for control households appears to be insignificant and modest at 0.8 percent (*p-value* $= 0.36$).

4.5.2 Difference-in-differences (DID) estimation results

Results from the DID regression models are presented in Table 11 and Table 12 for the stem and split-off households, respectively. In presenting the results, four model categories are included to shed light on the estimated effect of the main variable of interest as well as model performance as more and more control variables are included: 1) base model, 2) additional socioeconomic (excluding income), 3) socioeconomic and income, and 4) inclusion of the built environment indicators.

For the stem households, the base model (1), as shown in Table 11, indicates that exposure to the urban environment might reduce transportation expenditure by approximately 29.23 percent relative to the households that remained rural, all else equal. As additional control variables are incorporated in the models, this estimated effect appears to become less pronounced whilst the goodness-of-fit of the model improved noticeably. For instance, model (3), where socioeconomic and income indicators are incorporated, suggests that relocating to urban environment could reduce the share of transportation expenditure from a total income by 11.83 percent relative to the control households, a fraction to the estimated 29.23 percent as shown in the base model (1). It should be noted, nonetheless, that none of the DID estimators is statistically significant.

Table 11 – Model results, stem households

	Dependent variable: Monthly household transportation expenditure from a total income (logged)			
	(1)	(2)	(3)	(4)
Treatment	0.090	-0.005	0.060	0.131
	(0.248)	(0.244)	(0.241)	(0.256)
Time	-0.016	-0.113	-0.057	0.007
	(0.248)	(0.249)	(0.245)	(0.255)
Treatment*Time (DID)	**-0.346**	**-0.313**	**-0.126**	**-0.199**
	(0.350)	(0.342)	(0.345)	(0.375)
Income (logged)			-0.274**	-0.296**
			(0.118)	(0.121)
Socioeconomic	N	Y	Y	Y
Built Environment	N	N	N	Y
Observations	120	120	120	120
R-squared	0.020	0.109	0.151	0.161
Adjusted R-squared	-0.005	0.045	0.081	0.067
Note:				*p<0.1; **p<0.05; ***p<0.01

Note: For brevity, socioeconomic and built environment covariates are not presented.

In terms of goodness-of-fit of the models associated with the inclusion of additional control variables, it is apparent that incorporating socioeconomic variables, particularly income, exerts a meaningful effect to improve model performance. As Table 11 indicates, the explanatory power of the model improves considerably relative to the base model as control variables were included. However, despite the improved model performance associated with the inclusion of control variables, the absolute explanatory power (R-squared) of the models itself is reasonably low as none of the models estimated could explain more than 20 percent variance of the dependent variable.

For the split-off households, as shown in Table 12, the base model (1) suggests that relocating to urban areas might reduce transportation expenditure from a total income by approximately 17.03 percent (*p-value* < 0.1) in comparison to the control households, holding other variables constant. When considering the inclusion of socioeconomic and income indicators as shown in model (3), the model suggests a statistically insignificant estimate of 10.62 percent reduction in the share of household transportation expenditure from a total income, all else equal. Notwithstanding the insignificance of the coefficient, this value appears to be relatively similar to the one observed in the model for stem households (i.e., 11.83 percent).

Table 12 – Model results, split-off households

	Dependent variable: Monthly household transportation expenditure from a total income (logged)			
	(1)	(2)	(3)	(4)
Treatment	0.206**	0.211**	0.226***	0.193**
	(0.091)	(0.090)	(0.082)	(0.084)
Time	-0.059	-0.005	0.144	0.140
	(0.091)	(0.098)	(0.090)	(0.090)
Treatment*Time (DID)	**-0.187***	**-0.198**	**-0.112**	**-0.060**
	(0.128)	(0.128)	(0.117)	(0.133)
Income (logged)			-0.494***	-0.497***
			(0.035)	(0.035)
Socioeconomic	N	Y	Y	Y
Built Environment	N	N	N	Y
Observations	1,052	1,052	1,052	1,052
R-squared	0.010	0.035	0.192	0.198
Adjusted R-squared	0.007	0.027	0.185	0.189
Note:				*p<0.1; **p<0.05; ***p<0.01

Note: For brevity, socioeconomic and built environment covariates are not presented.

Also similar to the behavior of the models as observed in estimation for stem households, it appears that the goodness-of-fit of the models for split-off households is relatively low across the models estimated. As indicated, none of the models could explain more than 20 percent variance of the dependent variable.

The considerably low predictive power indicates that there are unobserved factors influencing household transportation expenditure. These unobserved factors are likely difficult to measure and therefore present a challenge to be incorporated in the model. While largely speculative, these factors may include aspects such as distinct social networks between rural and urban households that allow for a divergent on how the households in each geographic location fulfill transportation needs. Uncovering or shedding further light on these factors likely require a qualitative assessment, which is beyond the scope of this dissertation.

4.5.3 Counterfactual graphs

The application of the DID approach allows an evaluation of counterfactual scenarios, or the possible circumstances had the treatment households did not relocate and instead followed the trajectory of the control households. These counterfactual scenarios, derived from the regression coefficients, are shown in Figure 22 and Figure 23, representing the analyses for the stem and split-off households, respectively.

As indicated in both graphs, it appears that the treatment households deviate from the counterfactual trend derived from the trajectory of control households. This deviation reflects the reduction of transportation expenditure from a total income that the treatment households realized as they relocated to an urban environment, which is observed in both stem and split-off households.

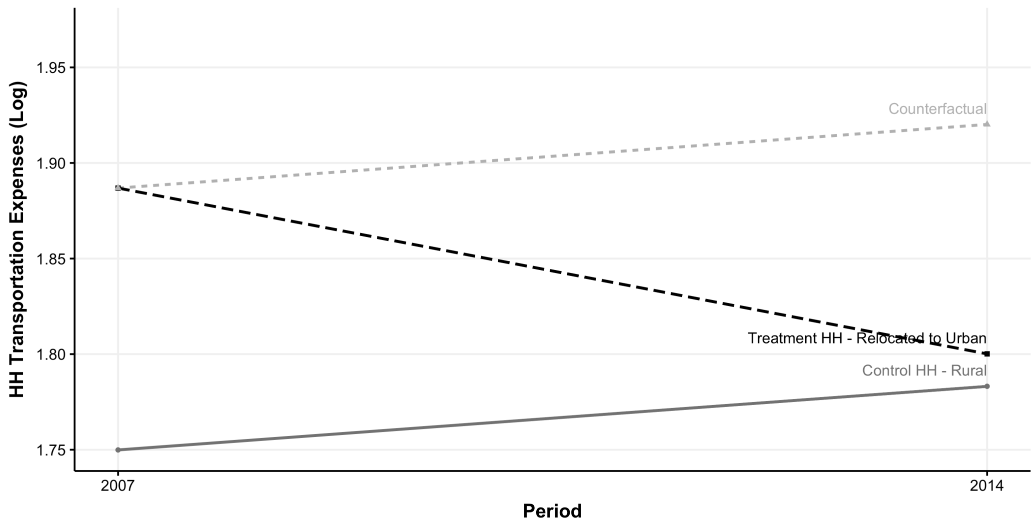

Figure 22 – Counterfactual graph indicating the possible scenarios had treatment households did not relocate to urban areas, stem households

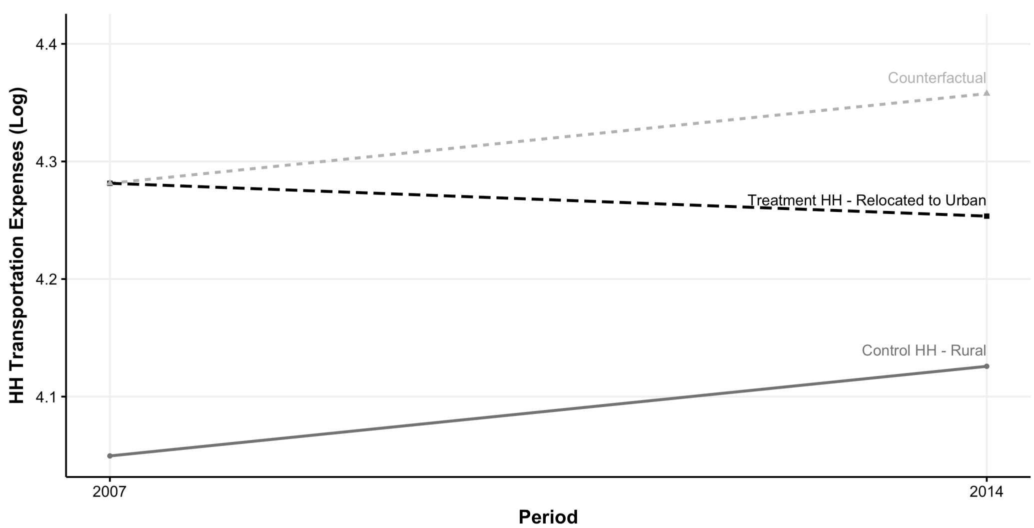

Figure 23 – Counterfactual graph indicating the possible scenarios had treatment households did not relocate to urban areas, split-off households

4.6 Conclusion

In this chapter, a series of analyses are presented that revolves around estimating the effects of relocating to urban areas on a travel behavior outcome, i.e., the share of household transportation expenditure from a total income. By leveraging the longitudinal character of the IFLS and exploiting its prominent feature that tracks internal migration of the sampled households over time, this study proposes a before-after, treatment-control evaluation to causally estimate the impact of relocating to urban areas on household transportation expenditure). In that regard, relocating households who were rural in 2007 and became urban in 2014 are considered as treatment households. Households who were rural in 2007 and remained rural or relocated to other rural areas are assigned into a control group. This two-by-two factorial that captures before-after, treatment-control evaluation allows causal estimation using DID models.

Results indicate, albeit statistically insignificant, that relocating to urban areas could reduce the share of household transportation expenditure by approximately 11 percent. These seemingly modest results might not be particularly surprising considering the findings in the literature of travel behavior and built environment interaction.

Nonetheless, for planners and policymakers, the findings highlight the relative importance of ensuring the supply of compact, dense, and spatially mixed urban environment, particularly in places where relocating households tend to locate. Given that rural-urban migration households might be able to realize transportation expenditure-reduction benefits associated with exposure to a compact urban environment, they could instead allocate resources into a range of utility-maximizing consumptions, such as for food, clothes, savings, to name a few.

Several built environment support systems and planning tools might be considered to achieve the policy initiatives on advancing the development of compact and spatially mixed urban environment. For one, provisions of affordable housing in or near the city center or regional employment clusters aimed to provide relocating households with residential options might offer a viable policy measure, following Guerra's (2013) suggestion based on a study of Mexico City. Specific to the Indonesian context, it also speaks to the importance of integrating relocating households with the existing fabric of *kampung*, which characterizes a large proportion of Indonesia's urban landscape (Monkkonen, 2013; Setiawan, 2010). In addition, a traditional planning tool in the form of land use code, or better known as RDTR (*Rencana Detail Tata Ruang*) and RTRW (*Rencana Tata Ruang Wilayah*) in Indonesia, remain of importance to achieve the key policy objectives.

Examples of policy measures, as mentioned above, seek to mitigate the propensity of rural-urban households to reside in distant places from the city center and or employment clusters, which typically have low accessibility to amenities and job opportunities. Under the condition where rural-urban migrant households could only afford or, for lack of better words, forced to choose to live in inaccessible and peripheral neighborhoods characterized by a leapfrog development pattern, the relocating households might not be able to realize transportation expenses-reduction benefits due to, for example, excessive commuting.

Taken together, this study contributes to the travel behavior literature on several fronts, as seen from the perspective of theoretical, methodological, and empirical contributions.

Theoretical contribution. This study contributes to advancing the dialogue on how life events, particularly residential relocation in the form of rural-urban migration, could shape travel behavior. As the qualitative observation of the data and DID estimation results of this study indicate, rural-urban migration appears to induce travel behavior changes, as shown in the reduction of household transportation expenditure, as a proxy for travel demand. This study also contributes to developing a theoretical framework, as shown in Figure 13. This framework can be adopted into a replicable research design for future studies concerning the interrelated association between residential relocation and travel behavior.

Methodological contribution. This study is among the few, if not the first, that constructs a before-after, treatment-control approach using longitudinal data separated over several years of span to study the effects of residential relocation. Previous studies that also use a before-after treatment-control approach using household or individual-level microdata tend to focus on travel behavior dynamics in a relatively short period, e.g., before-after evaluation of Expo LRT in Los Angeles (Spears et al., 2016). This study also illustrates the application of difference-in-differences (DID), as an example of causal estimation methods, to conducting travel behavior research. To date, the application of causal inferences in travel behavior research remains noticeably limited (Brathwaite & Walker, 2018; Coevering et al., 2016; Næss et al., 2018), which likely stems from the prevalence of cross-sectional study in the prevailing studies.

Empirical contribution. From an empirical standpoint, this study further advances empirical research on travel behavior in developing countries. In light of the notion that the large majority of travel behavior studies have so far disproportionately focused on empirical cases from developed regions, this study provides a refreshed perspective by assessing factors associated with travel behavior in Indonesia as an example of the world's emerging economies.

Despite the contributions this study could make, there are several limitations. For instance, the increased sample size could likely provide stronger statistical estimations, and future studies might as well incorporate a greater number of respondents. Despite this apparent limitation, the herculean efforts that the IFLS administrator had put in place to track the relocating households across Indonesia where infrastructure are appreciated – not to mention the challenges they had to endure in conducting a longitudinal survey in a country where the technological means is not as advanced as in developed countries (Thomas et al., 2001, 2012). Another limitation is that, at this moment and to the extent of the literature, there are no comparable studies on travel behavior effects of rural to urban residential relocation using a before-after treatment-control evaluation as applied in this study. This proposition limits the capacity to situate the findings in the literature. Furthermore, third, this study only estimates internal migration given the data availability. Relocating households that moved outside of the country are not captured.

These limitations should be seen as a concrete opportunity to pave the avenue for future research. For instance, a research endeavor on the travel behavior effects associated with residential relocation from suburban to urban, and maybe vice versa, might as well be of interest to researchers and policymakers. In light of the rapid urbanization occurring in multiple corners of the world, the need to empirically assess and evaluate how urbanization, in the form of rural-urban migration, impacts travel behavior is of critical importance.

CHAPTER 5. TRAVEL BEHAVIOR OF URBAN NON-MOVERS FROM A PANEL PERSPECTIVE

5.1 Introduction

In the preceding chapter, the focus was on travel behavior effects of rural to urban migration in urbanizing Indonesia. As a natural segue and with respect to the overall research framework, this chapter then examines the dynamic aspects of travel behavior for urban non-movers as the built environment and socioeconomic traits evolve. From the lens of built environment changes, the urbanization process in Indonesia entails that certain urban regions were becoming denser and hosting a more diverse array of land uses and amenities. From a socioeconomic standpoint, as Indonesia is urbanizing and due to the household and individual life-course and progression, the socioeconomic traits of the country's populace are changing as well. Considering the theoretical framework of travel behavior that suggests the built environment and socioeconomic as the two primary determinants, this chapter aims to address the following question: How do the built environment and socioeconomic changes influence travel behavior over time for non-mover households?

This chapter is structured as follows. The following section elaborates on the extent of the literature on panel data analyses of non-movers. An examination of the literature on this subject mirrors the general literature review, as presented in Chapter 2, and suggests the noticeable lack of studies that specifically focus on the non-movers from a longitudinal perspective. Following the literature review, the subsequent sections further discuss the conceptual framework, data structure, empirical strategy, and the estimation results. Potential planning implications informed by the findings, as well as contributions of this particular chapter, are discussed in the conclusion section.

5.2 Literature Review: Panel Data Analyses of Non-Movers

A review of the literature suggests a noticeably small number of studies on panel data analyses of non-movers. Most studies using panel data typically include an indicator of residential relocation, which does not necessarily conform to the primary aim of this chapter. It is also apparent that the majority of the studies in this strand of literature use multipurpose household panel surveys, which is likely a testament to the lack of conventional travel surveys that track the respondents over time.

The extensive use of multipurpose household panel survey also explains why the existing studies typically model an indicator that is widely available in such surveys, i.e., vehicle ownership. For instance, Prillwitz, Harms, & Lanzendorf (2006) use panel data of the German Socioeconomic Panel (GSOEP) to analyze the likelihood of a given household to increase or decrease car ownership. They estimate binomial probit models and discover variables such as the birth of the first child and rising income are strong predictors of an increase in car ownership. A study by Clark et al. (2014) leverages the "true" longitudinal characteristic of UK Household Longitudinal Study (UKHLS) to explore the relationship between life-course events and travel behavior, where car ownership is the primary dependent variable. They develop models to estimate factors associated with the increase and decrease in household car ownership. Their findings indicate the varying extents of residential relocation, employment changes, and household structure to influence car ownership decision making.

Also using UKHLS data, a 2016 study separately estimates the factors influencing changes in car ownership level (i.e., from zero to 1 vehicle, 1-2, 2-1, 1-0) (Clark et al., 2016). Despite using panel data, they structure the estimation approaches based on a cross-sectional assumption where a variety of baseline observations (e.g., neighborhood context in the previous survey wave) are regressed to estimate the changes in car ownership level.

While illuminating, the two studies (Clark et al., 2014, 2016) incorporate spatial indicators simply at baseline observation and somewhat overlook its dynamics over time. They argue, "…that spatial context variables are unlikely to change significantly between two consecutive waves" (Clark et al., 2016, p. 573), which might be appropriate given the short gap between the waves used in the data. However, this hypothesis remains unexplored using panel data separated over a long-time period where substantial changes in spatial context might be observed and therefore remains largely unexplored.

Yamamoto (2008) conducts comparative analyses of factors associated with car ownership using France and Japanese longitudinal datasets. The French dataset originates from a "true" panel survey while the Japanese one comes from a retrospective survey. Yamamoto (2008) finds more or less similar predictors of car ownership between these two cases, particularly the effects of the number of adults, despite substantial differences in the modeling approaches between these two datasets.

Several studies on car ownership using panel data, as discussed above, seem to indicate the conventional approach of modifying the dependent and explanatory variables into a numerical change of the value across survey periods (Clark et al., 2014, 2016; Prillwitz et al., 2006; Yamamoto, 2008). The commonly used methods to account for the changes in vehicle ownership and associated explanatory variables are logit and probit regression.

In addition to the strand of literature on vehicle ownership from the panel perspective as discussed above, a body of literature on panel data analyses of vehicle use and mobility pattern has also been identified. Scheiner & Holz-Rau (2013a) use German Mobility Panel data to assess annual changes in the number of trips a given individual made per day from data of approximately 6,900 individuals. Using cluster-robust regression estimates instead of random effects model since this model "assumes constant correlation between successive observations of the same unit" (Scheiner & Holz-Rau, 2013a, p. 170), they discover noticeable effects of baseline indicators (e.g., municipality sizes and central city residents) as predictors of annual changes in the number of trips by mode.

Wasfi et al. (2017) examine factors associated with walking behavior from a panel perspective using the Neighborhood Population Health Survey of Canada (1994 – 2010), which was collected biennially. The dependent variable represents a binary indicator of the walking behavior of a given individual. The primary variables of interest are age and length of exposure to a variety of neighborhood walkability characteristics. They treat exposure as a time-invariant variable while age and a host of other covariates as time-variant variables. Using mixed-effects logistic regression, the authors find relative influences of different neighborhood characteristics on walking behavior mediated by the age of the respondents.

A study by Coevering et al. (2016) is probably the closest match of this intended research on panel data analyses of non-movers from a long-term perspective. Instead of observing travel behavior dynamics within a reasonably short period (approximately 1-year gap) as commonly used in studies as described above, they assess the dynamics over a long period. Their focus is not necessarily on residential relocation itself, despite a small subset of their samples (250 out of 1322 observations, or 19%) relocated to new residential relocation over the course of 7 years. Coevering et al. (2016) employ a cross-lagged panel model approach to explore a more robust causal interpretation of travel behavior and built environment relationship.

Existing studies of travel outcomes from a panel perspective, as discussed, point to more diverse methodological approaches than studies of vehicle ownership. While the literature on vehicle ownership appears to somewhat uniformly revolve around the application of logit and probit modeling approaches, the methods in the studies of vehicle use and travel outcomes show more variation.

All in all, concerning the inquiry of panel analyses of non-movers, the extent of the literature review points to the notion that no study specifically assesses this inquiry. Virtually every study incorporates indicators related to residential relocation. In addition, only a very few reviewed studies take into account travel behavior dynamics over a reasonably long-time period. Most studies mainly rely on approximately less than three years of panel observations. These factors point to the gap in the literature and motivate this inquiry, which focuses on the travel behavior dynamics associated with the evolving spatial context of urban non-movers over a considerably long-time period.

5.3 Conceptual Framework and Hypotheses

5.3.1 Conceptual Framework

Figure 24 presents the conceptual framework of this chapter that revolves around analyzing factors associated with travel behavior dynamics from a panel of urban non-movers. As shown, the central contribution of the study, as presented in this chapter, stems from incorporating indicators over time to model travel behavior, particularly household transportation expenditure, since existing studies tend to model the association from a one-time approach due to its heavy reliance on cross-sectional observations.

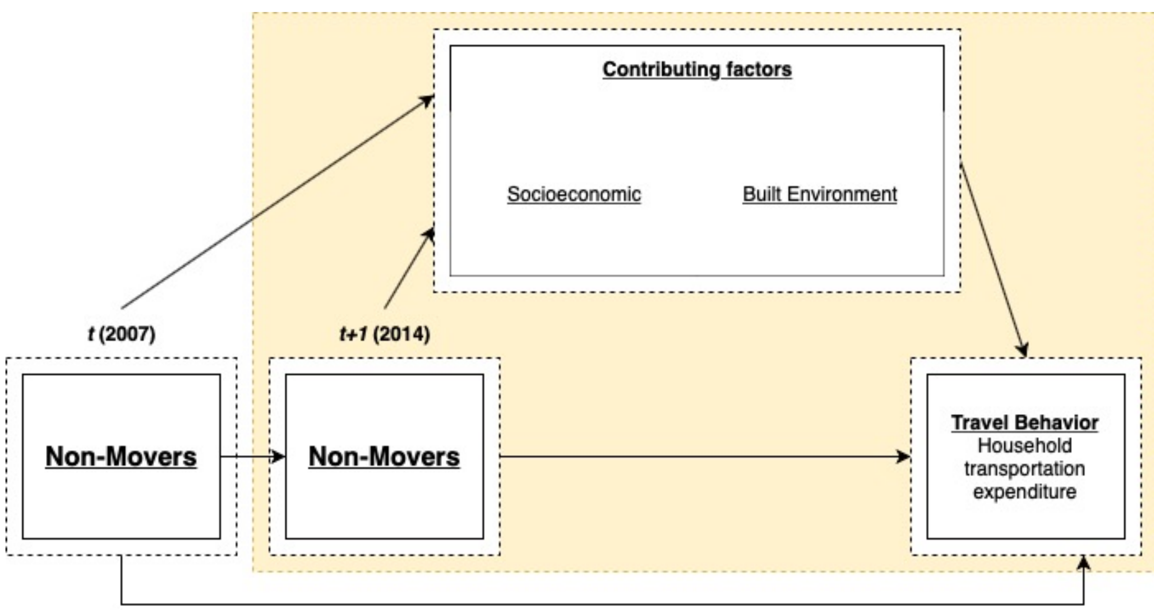

Figure 24 – Travel behavior as a function of panel analyses of non-movers

5.3.2 Hypotheses

Following the theoretical proposition on the determinants of travel behavior, present cross-sectional studies on the factors associated with household transportation spending (Guerra, 2017; Guerra et al., 2018), as well as considering the research framework presented in Figure 24, the main hypothesis expects that socioeconomic and built environment changes would exert significant influences on transportation expenditure over time. The alternative hypothesis, however, is the notion that travel behavior is relatively stable over time; hence the share of transportation expenditure would not change despite observable changes in the socioeconomic and built environment attributes. This proposition is particularly relevant to the built environment since it could take years to observe substantial changes in the physical spaces, even in urbanizing Indonesia; therefore, the influences of the built environment, if any, are likely to be modest.

Q2: For urban non-movers, how do the dynamic of the built environment and socioeconomic changes influence travel behavior?

H_{Main}: Built environment and socioeconomic changes influence non-mover households' share of transportation-related expenses.

$H_{Alternative}$: Travel behavior is relatively stable over time. Therefore, changes in the built environment or socioeconomic traits do not influence non-mover households' share of transportation-related expenses.

5.4 Data and Methodology

5.4.1 Data

Similar to the primary data source for addressing the analytical chapter on rural-urban migration, this chapter also uses IFLS 4 (2007) and IFLS 5 (2014) to construct the dataset and address the research question on the dynamic of travel behaviors for urban non-movers. The dependent variable of interest in this analytical chapter of urban non-movers is also the same as estimated in the rural to urban migration chapter, i.e., the share of transportation-related expenses from a total income.

Concerning the 7-year gap between the two waves of IFLS data, at least two factors might explain the relative appropriateness of using that timeframe to observe Indonesia's urbanization process and how does it relate to travel behavior changes for urban non-movers. Firstly, considering that travel behavior literature indicates that the influences of the built environment on travel behavior appear to be modest, as indicated in the prior sections, the quest to observe travel behavior dynamics would, therefore, be more appropriate, theoretically, in places that experienced relatively rapid built environment changes. This proposition lends a hand, again, to signify why Indonesia is an appropriate case since the change in the average density of the country's urbanized areas over the 10-year period (2000-2010) is among the most dramatic in comparison to other countries in the EAP region (World Bank, 2015). Moreover, while the IFLS dataset used in this chapter does not precisely coincide with the time frame in World Bank's (2015) study, the data derived from IFLS 4 (2007) and 5 (2014) might still capture that urbanization and built environment dynamics. Secondly, the appropriateness of a 7-year gap to observe travel behavior dynamics is somewhat further validated by the proposition in an earlier study by Clark et al. (2016). In this study, they use two waves of UKHLS data separated over 12 months to study factors associated with the change in vehicle ownership level. However, instead of pooling the built environment indicators from each survey wave to model vehicle ownership, the authors opted to only incorporate spatial contexts at the baseline condition, or the first wave of the UKHLS in 2009, since they expect the built environment does not evolve substantially over a 12-month period.

Considering the factors as presented above, this following identification strategy is therefore applied to construct the dataset:

a. Select households that are coded as '96' in the residential move column (sc21x) in IFLS 5 (module bk_sc), which indicates that the respective households did not relocate between 2007 and 2014.
b. Conditional on *a*, select urban households as coded as '1' in the binary indicator of urban and rural designation column (sc05, module bk_sc) in both IFLS 4 (2007) and IFLS 5 (2014).
c. Link the geocoded information of the sampled households from IFLS with the Village Census (*Potensi Desa* – PODES) data and select the samples that meet these following requirements to ascertain more or less similar urbanization characteristics across the sampled households: 1) Population density[5] at the sub-district level is more than > 1,000 people/km^2 and 2) Road condition for all villages within a given sub-district is considered good, can be traversed through all year, and covered by asphalt instead of pebble or simply plain soil[6].

Based on this identification strategy, the assembled socioeconomic indicators from several IFLS 4 and IFLS 5 modules, as well as built environment covariates derived primarily from Village Census data that made up the balanced panel dataset of 1128 sampled households are listed in Table 13.

[5] The decision to subset the sample using these 1,000 people/km2 threshold stems from the notion that this value is the threshold with respect to population density at village level used by the *Badan Pusat Statistik* (2010), Government of Indonesia to estimate rural-urban village classification.
[6] The indicator on road condition is derived from these following questions as documented in the Village Census surveys: "Types of road surface of the widest road: Asphalt/Concrete – 1; Pebble – 2; Land – 3; Others: - 4" and "Is it can be passed through by four wheel vehicle all along year? Yes – 1; No – 2".

Dependent variable. Similar to the first analytical chapter on rural-urban migration, the dependent variable in this chapter is the share of household transportation expenditure from a total household income, which serves as a proxy for travel demand.

Table 13 – Urban non-movers: List of variables

Variable	Description	Source	
		IFLS 4 Module	IFLS 5 Module
Dependent variable			
Transportation expenditure	Share of transportation expenditure from total income	b1_ks2	b1_ks2
Covariates (Socioeconomic)			
Income	Household monthly income[1] (Indonesian Rupiahs - IDR)	bk_ar1; b3a_tk2; b3a_re	bk_ar1; b3a_tk2; b3a_re
Age	Average age of the household	bk_ar1	bk_ar1
Household (HH) size	Total number of household members	bk_ar1	bk_ar1
Employed HH members	Number of employed household members	bk_ar1	bk_ar1
HH members (< 5 years old)	Number of household members aged less than 5 years old	bk_ar1	bk_ar1
HH members (> 65 years old)	Number of household members aged 65 years old or older	bk_ar1	bk_ar1
Vehicle ownership	Number of household members who own vehicle	b2_hr1	b2_hr1
Covariates (Built environment)			
Population density	Number of population per square kilometer	Village Census, 2008 & 2011	
Household density	Number of household per square kilometer	Village Census, 2008 & 2011	
Education facilities density	Number of kindergartens, elementary schools, junior high schools, and high schools per square kilometer	Village Census, 2008 & 2014	
Retail density	Number of retail establishments[2] per square kilometer	Village Census, 2008 & 2014	
Distance to city center	Euclidean distance from the respondent's neighborhood centroid to the neighborhood with highest retail density in the corresponding district	Author's calculation using 'geosphere' package in R Studio	

Note: [1] Household monthly income is comprised of primary/regular income as well as secondary and a retirement pension, as applicable depending on the household.

[2] Retail establishments represent the sum of minimarkets, restaurant/food stalls, food and beverage stores/*warung*, grocery stores/*kelontong*, and hotels and inns/motel in a given neighborhood

Control variables. Following the literature of travel and built environment, the socioeconomic covariates included in the dataset primarily serve as control variables. While these socioeconomic covariates are considered as control variables, the literature suggests that socioeconomic indicators are the main predictors of travel behavior.

Main variables of interest. The built environment indicators included in the dataset are the main variables of interest to be evaluated. These indicators include variables commonly incorporated in the literature, e.g., household density, retail density, distance to the city center. Not all built environment variables listed in Table 13 might be incorporated in the regression model as a means to minimize the multicollinearity issue associated with the notion that places with high population density likely maintain, say, a considerable presence of educational facilities.

5.4.2 Sample Characteristics

Table 14 presents the sample characteristics with an emphasis on highlighting the dynamics or changes of the variables incorporated in this study between 2007 and 2014. In terms of the share of transportation expenditure from a total income as the dependent variable in this study, it appears there was a 1.19% reduction (Figure 25). A paired t-test analysis suggests that this reduction is statistically significant at the 99% level (*p-value* = 0.002).

Other indicators, as presented in Table 14, also point to the statistically significant changes between 2007 and 2014 associated with the life progression of the sampled households. For instance, the data suggest a substantial change of inflation-adjusted income where an average increase of 1.36 million IDR (or approximately US$100 as of January 2020 exchange rate) is observed (*p-value* < 0.01). While average income increased, the average number of employed household members in the sampled households actually decreased from 1.68 in 2007 to 1.56 in 2014 *(p-value* < 0.01).

Table 14 – Urban non-movers: Descriptive statistics of the sampled households (n=1128), 2007-2014

Variable		Descriptive Statistics				Mean diff.	*p*
		Mean	S.D.	Min	Max		
Dependent var.							
Transportation expenditure	07	0.10	0.10	0.01	0.55	-0.01	***
	14	0.09	0.09	0.01	0.49		
Socioeconomic							
Income (IDR)	07	3,186,549.0	3,064,109.7	105,328.2	28,288,150	1,360,914.6	***
	14	4,547,463.6	3,816,115.1	100,000.0	37,083,333		
Age	07	28.86	11.88	7.33	81.00	3.30	***
	14	32.16	13.06	7.67	78.00		
Household (HH) size	07	4.02	1.86	1	14	-0.13	***
	14	3.89	1.74	1	14		
Employed HH members	07	1.68	1.18	0	7	-0.12	***
	14	1.56	1.06	0	6		
HH members (age < 5)	07	0.55	0.65	0	3	-0.13	***
	14	0.42	0.64	0	4		
HH members (age > 65)	07	0.19	0.45	0	2	0.02	
	14	0.21	0.49	0	3		
Vehicle ownership	07	1.42	0.98	0	7	-0.05	
	14	1.37	0.90	0	5		
Built environment							
Population density	07	7,301.21	7,447.25	1,000.68	49,934.77	626.28	***
	14	7,926.49	7,190.25	1,069.33	48,940.08		
Household density	07	1,881.71	1,875.41	255.34	12,587.63	230.54	***
	14	2,112.25	1,906.54	283.35	11,698.83		
Retail density	07	93.46	101.39	4.03	938.25	16.69	***
	14	110.16	112.75	3.32	1,174.96		
Education facilities	07	7.43	6.44	1.12	33.96	0.35	***
	14	7.78	6.27	0.79	33.96		
Retail-HH Balance	07	0.06	0.03	0.01	0.21	0.00	
	14	0.06	0.03	0.01	0.40		
Distance to city center	07	5.79	6.30	0.00	33.77		
	14	5.79	6.30	0.00	33.77		
Note:						*p<0.1; **p<0.05; ***p<0.01	

In addition, the data also capture a slight reduction in household size (*p-value* = 0.007) and the number of household members under five years old (*p-value* < 0.01). As might be expected, the average age of the sampled households increased from 28.86 to 32.16 years old *(p-value* < 0.01) between 2007 and 2014. As average age increased, the average number of household members aged 65 years old or more increased from 0.19 to 0.21, despite a paired t-test analysis that suggests that this change is statistically insignificant (*p-value* = 0.126).

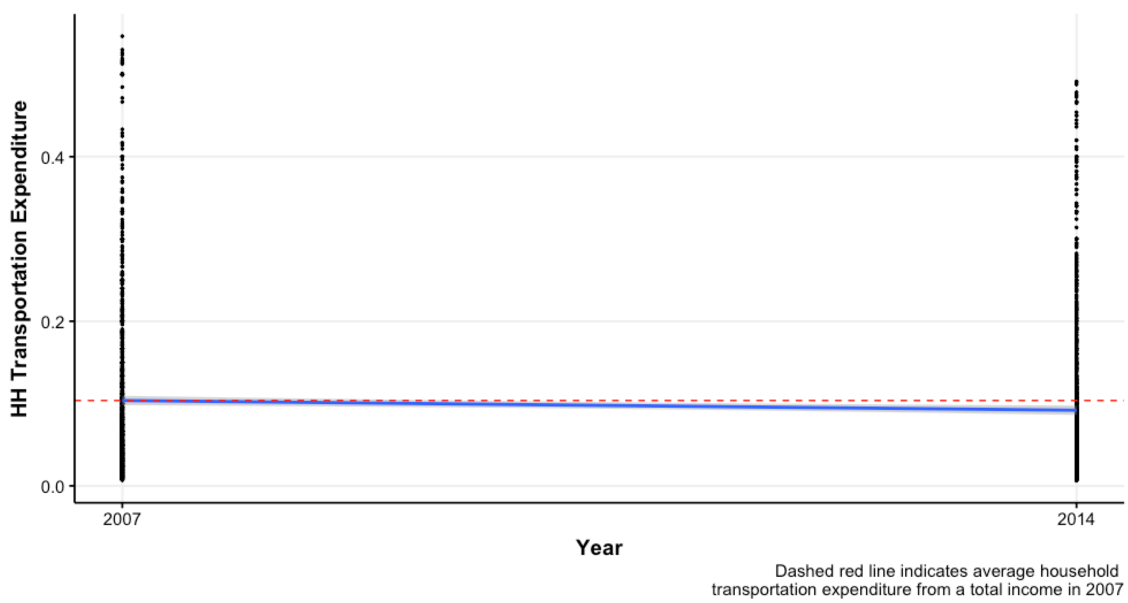

Figure 25 – Average change in household transportation expenditure from a total income, 2007-2014

Concerning the dynamics of the built environment, Table 14 seems to substantiate the urbanization phenomenon occurring in Indonesia. As can be seen, on average, population and household density increased by 625.28 people/km2 and 230.54 households/km2, respectively. A paired t-test for each of these indicators suggests that the changes are both statistically significant at the 99% level. Given that the data suggest that on average household size had decreased (Table 14), it might be inferred that the increase in population and household density within the sampled neighborhoods in the dataset is likely driven by the influx of new population. This influx, therefore, likely induced the proliferation of new business establishments as retail density increased by 16.69 establishments/km2 from, on average, 93.46 establishments/km2 to 110.16 establishments/km2 between 2007 and 2014 (*p-value* < 0.01).

Figure 26 shows the geographical distribution of the sampled households at the district/*kabupaten*-level. As can be seen, a substantial number of the sampled households resided in the Java island.

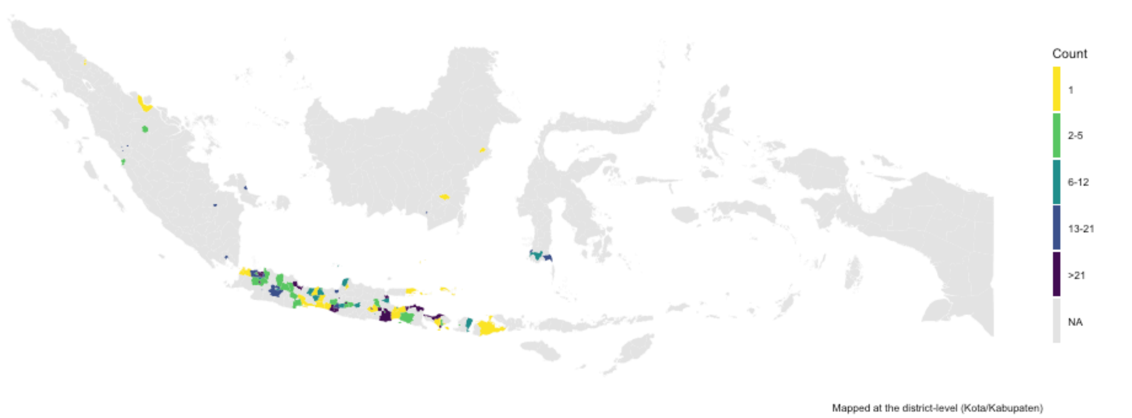

Figure 26 – Geographic distribution of urban non-mover households (Mapped at the district-level)

Table 15 further illustrates the geographic distribution shown in Figure 26 by indicating in both absolute values and percentage terms of the sampled households' residential location by the province as well as the percentage distribution by region. As shown in the table, provinces in Java host 77.13 percent of the respondents, followed by Sumatera region (8.95 percent), Bali and Nusa Tenggara (8.24 percent), Sulawesi (3.72 percent), and Kalimantan (1.95 percent).

Table 15 – Provincial distribution of urban non-mover households

Province	Region	2007-2014		
		# Households	%	% (by region)
Bali	Bali & NT	75	6.65	8.24
West Nusa Tenggara	Bali & NT	18	1.60	
Banten	Java	22	1.95	77.13
Central Java	Java	147	13.03	
East Java	Java	158	14.01	
Jakarta	Java	219	19.41	
West Java	Java	142	12.59	
Yogyakarta	Java	182	16.13	
South Kalimantan	Kalimantan	21	1.86	1.95
East Kalimantan	Kalimantan	1	0.09	
South Sulawesi	Sulawesi	42	3.72	3.72
Bangka Belitung Island	Sumatera	15	1.33	8.95
Lampung	Sumatera	19	1.68	
North Sumatera	Sumatera	1	0.09	
Riau	Sumatera	4	0.35	
South Sumatera	Sumatera	13	1.15	
West Sumatera	Sumatera	49	4.34	
	Total	1128	100	

In addition to breaking down the distribution of the sampled households at the provincial-level, further analyses were conducted to indicate the distribution of sampled households by the district's (*Kota* or *Kabupaten*) urbanization characteristics following the typology as specified by the World Bank (2018a). The typology centers around classifying whether a given district is categorized as multi-district metro, single-district metro, urban non-metro. Due to the substantial variation of districts in terms of urbanization rate and development stage within the classification of multi-district metro, a slight modification from the original typology is made and therefore several districts are grouped separately, i.e., Jabodetabek and Surabaya are grouped as one group while Bandung, Medan, and Yogyakarta in another. Subsequently, by linking the districts of the sampled households with this modified typology derived from the World Bank (2018a), the break down is given as follows: multi-district metro (286 households, 25.3 percent), single-district metro (38 households, 3.4 percent), urban non-metro (297 households, 26.3 percent), Bandung, Medan, and Yogyakarta (174 households, 15.4 percent), and Jabodetabek and Surabaya (328 households, 29.1 percent).

5.4.3 Estimation Approaches

Having developed the balanced panel dataset, a series of panel regressions are evaluated to estimate the factors associated with the dynamics of the share of transportation expenditure from a total income between 2007 and 2014. In evaluating panel regression techniques, the focus is to test whether random or fixed effects would be the appropriate option by conducting the comparative analyses using the ‘plm’ package in R Studio platform (Croissant et al., 2017).

The primary difference between the fixed- and random effects approach rests on the assumption between the two. In fixed effects model, it is assumed that the subject (or in this study, household) differences are accounted for; meanwhile, in random effects model, it is assumed that since the households are typically selected as random to participate in the panel data collection, their characteristics in question are expected to be random (Colonescu, 2016). In a practical manner, the fixed effects model centered around the differences within-subject can only estimate time-varying variables. On the other hand, the random effects model is able to estimate both time-varying and time-invariant covariates.

Theoretically, fixed effects would likely be a more appropriate option than random effects owing to the notion that most variables of interest in this study are time-variant.

Only one indicator incorporated in the dataset is a time-invariant variable, i.e., distance to the city center. Following several estimations approaches in previous studies (Colonescu, 2016; Croissant et al., 2017; Parady et al., 2014), the fixed effects estimator of this study can be represented in this following equation.

$$y_{it} = \alpha_i + \beta_{it}SE + \beta_{it}BE + e_{it} \tag{5}$$

where α represents unobserved heterogeneity at the household level.

In this study, t_{2007} represents the time in 2007 and t_{2014} represents the time in 2014. Considering t_{2007} and t_{2014}, the equation for each time for the i household would be:

$$y_{i2007} = \alpha_i + \beta_{i2007} + \beta_{i2007}SE + \beta_{i2007}BE + e_{i2007} \tag{6}$$

$$y_{i2014} = \alpha_i + \beta_{i2014} + \beta_{i2014}SE + \beta_{i2014}BE + e_{i2014} \tag{7}$$

where y_i represents household's travel behavior, $\beta_i SE$ and $\beta_i BE$ represent time-variant socioeconomic and built environment changes, respectively, and e_i is the error term. From these equations, the fixed-effects model transforms the data comprised of observations from two or more periods by subtracting the mean value over time (by the group) to the variables provided in the model (Croissant et al., 2017; Parady et al., 2014). This process produces a *time-demeaning* equation, as shown below:

$$y_t - \bar{y} = \alpha - \bar{\alpha} + \beta(SE_t - \overline{SE}) + \beta(BE_t - \overline{BE}) + + e_t - \bar{e} \tag{8}$$

The fixed effects model, as presented in the equations above, solely incorporate the individual or entity effect. The model can, therefore, be extended to account for both the individual and time-fixed effect by including the time indicator (μ_t), as shown in the equation below (Croissant et al., 2017; Endsley, 2016).

$$y_{it} = \alpha_i + \beta_{it}SE + \beta_{it}BE + \mu_t + e_{it} \tag{9}$$

The random effects approach subscribes to the notion that "since the individuals in the panel are randomly selected, their characteristics,… should also be random" (Colonescu, 2016). The primary differentiating characteristic between random and fixed effects model lies in the presence of individual-specific random term (u_i), which in addition to the "initial regression error term" (e_{it}), makes up the random effects error term (v_{it}) (Colonescu, 2016). The random effects equation in this study can, therefore, be represented by the following equation.

$$y_{it} = \alpha_i + \beta_{it}SE + \beta_{it}BE + v_{it} \tag{10}$$

Considering the fixed effects and random effects approach as presented, the following section discusses the results of estimating factors associated with household transportation expenditure using both approaches, followed by evaluating which model would be more appropriate using the Hausman test. Discussion and interpretation of the results from the selected estimation approach will be presented subsequently.

5.5 Results

5.5.1 Full model

Table 16 presents estimation results from both fixed and random effects regression models that incorporate the entire sampled households (n=1128). As indicated, not every built environment indicator computed to develop the dataset, as listed in Table 14, is included in both estimation approaches to minimize the multicollinearity issue. Regarding which regression approach would be more appropriate, results from the Hausman test, as shown in Table 17, indicate that the fixed effects model is the recommended estimation approach (*p-value* < 0.000 or well below the 0.05 threshold). To this end, the description of estimation results and the discussion that revolves around the findings, as presented below, will focus on the results from the fixed effects model instead of random effects.

The fixed effects estimation results, as shown in Table 16 point, to one particular aspect: Household transportation expenditure is primarily a function of socioeconomic indicators and that the average influences of built environment are likely insignificant and, at best, modest. Concerning socioeconomic indicators that exert statistically significant influences, the results suggest that income, employed household members, and the presence of child(ren) as the three influential variables. For every additional employed household member, a 7.4 percent increase in the share of household transportation expenditure from a total income is expected, holding other variables constant. This is likely attributed to the notion of increasing travel demand due to the mobility required to conduct employment or related tasks. Another statistically significant indicator, i.e., presence of child(ren), suggests that each additional child in the household likely reduces household transportation expenditure by 11.04 percent, all else equal. This finding could be interpreted from at least two potential scenarios: First, having more child could force a given household to direct expenses that otherwise would be used to cover transportation needs to child-related expenditure; or, second, an increasing number of children might reduce overall travel demand since the parents or adults within the household might spend more time taking care the child(ren) at home and therefore allocate fewer time and resources to travel. As expected, since the measure of household transportation expenditure is closely associated with income, this indicator wields a statistically significant influence. Specifically, a one percentage point increase in income could reduce household transportation expenditure by 0.58 percent, holding other variables constant. This finding could be interpreted as follows: In light of the increasing average income of the country's population as captured in the data, this result suggests that a given household might not necessarily decide to travel more despite having access to more substantial monetary resources.

Table 16 – Full model results: Urban non-movers, fixed and random effects

	Dependent variable: Monthly household transportation expenditure from a total income (logged)			
	Fixed effects (1)		**Random effects (2)**	
	Estimate	Std. Error	Estimate	Std. Error
Socioeconomic				
Household (HH) size	0.047	(0.029)	0.0003	(0.016)
# Employed household members	0.071**	(0.035)	-0.001	(0.023)
# HH members <5 years	-0.117**	(0.046)	-0.105***	(0.035)
# HH members >65 years	-0.001	(0.081)	0.012	(0.048)
Vehicle ownership	0.038	(0.029)	0.050**	(0.021)
Average age of HH members	0.0001	(0.005)	-0.003	(0.002)
Income (logged)	-0.583***	(0.047)	-0.339***	(0.026)
Built Environment				
Household density	-0.00004	(0.0001)	-0.00001	(0.000)
Retail density	0.0001	(0.0004)	0.0002	(0.0003)
Distance to city center			0.008*	(0.005)
Observations	2,256		2,256	
District/*kabupaten* dummy			Yes	
R-squared	0.130		0.148	
Adjusted R-squared	-0.755		0.114	
				*p<0.1; **p<0.05; ***p<0.01

Table 17 – Hausman test

Statistic (chisq)	p-value	Parameter	
40.135	0.0000	9	One model is inconsistent

Regarding the built environment, it appears that none of the built environment indicators had statistically significant influences on household transportation expenditure from a total income. Nonetheless, the coefficient sign of one of the built environment indicators (i.e., household density) incorporated in the fixed effects model appears to meet the expectation. That is, for every additional increase in household density, the share of household transportation expenditure from a total income might be reduced by approximately 0.0001 percent, holding other variables constant. The model suggests an apparently counterintuitive finding as higher retail density is associated with increasing transportation expenses. This likely indicates the possible scenario where the presence of retail establishments in the given neighborhood could induce more trips, thereby increases household transportation spending.

5.5.2 Sub-group models

The results of fixed effects full model incorporating the entire sampled households (n=1128), as shown above, might mask the dynamics of travel behavior and built environment relationships at distinct regional characteristics. This is mainly due to the primary shortcoming of fixed effects where the time-invariant indicator (e.g., district/*kota* or *kabupaten* dummy) cannot be incorporated in the estimation. To this end, the subgroup models presented in this section focus on estimating the fixed effects regression by district urbanization typology. This includes regression models for the subset of sampled households residing in *urban non-metro*, *single-district metro*, *multi-district metro*, Jabodetabek and Surabaya, and Bandung, Medan, and Yogyakarta, as shown in Table 18.

Overall, the results seem to indicate that regional contexts and urbanization characteristics might play a role in influencing the relationship between travel behavior and the built environment. As indicated in Table 18, the indicator of household density appears to have both negative and positive relationships, albeit largely insignificant in most cases, with the household transportation expenses depending on the models.

Table 18 – Subgroup models: Fixed effects

	Dependent variable: Monthly household transportation expenditure from a total income (logged)				
	Urban non-metro	Single-district metro	Multi-district metro	Jabodetabek & Surabaya	Bandung, Medan, & Yogyakarta
	(1)	(2)	(3)	(4)	(5)
Built Environment					
Household density	-0.001***	-0.0002	-0.00004	0.00002	-0.001
	(0.0005)	(0.0002)	(0.0001)	(0.0001)	(0.0004)
Retail density	-0.001	-0.0001	0.0001	0.0001	0.00004
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
Observations	594	76	572	666	348
R-squared	0.140	0.360	0.114	0.151	0.292
Adjusted R-squared	-0.778	-0.714	-0.833	-0.747	-0.499
					*p<0.1; **p<0.05; ***p<0.01

Note: Standard error in parentheses.

For brevity, socioeconomic indicators are not presented.

One particular noticeable finding is the one derived from the urban non-metro model that indicates the statistically significant influence at a 99% level of household density on household transportation expenditure (*p-value* = 0.008). That is, it suggests that each additional unit increase in household density might reduce household transportation expenditure by approximately 0.12 percent, all else equal. This finding contrasts with the one observed in the Jabodetabek & Surabaya model that instead indicates the positive, albeit insignificant, and considerably modest relationship between household density and household transportation expenditure.

Considering the estimated results from each subgroup model, the results seem to suggest that built environment measures appear to exert a more pronounced influence on travel behavior in a less urbanized region (i.e., urban non-metro) than in already developed ones (e.g., Jabodetabek and Surabaya, the two largest metros in Indonesia). One potential explanation is that the travel behavior of households living in highly developed regions might have been molded, for an extended period, by the relative ease of accessibility associated with the dense and widespread presence of destination opportunities and amenities that characterize most developed and highly urbanized regions. This, in turn, might render the neighborhood-level built environment changes largely ineffective to shape travel behavior. In contrast, households in lesser developed regions might not have benefitted from the abundant presence of amenities at the region-level; thus, changes in the built environment at the neighborhood-level could likely exert influence in shaping travel behavior.

5.6 Conclusion

The aim of this chapter was to study the determinants of household transportation expenditure for urban non-movers from a panel perspective. It tests the hypothesis that built environment and socioeconomic changes would significantly influence non-mover households' share of transportation-related expenses as a proxy for travel demand. In testing the hypothesis, this study uses a balanced panel dataset of urban non-mover households (n = 1,128) assembled from the IFLS 4 (2007) and 5 (2014). It estimates the built environment and socioeconomic determinants of transportation expenditure using a series of fixed effects panel regressions. Results indicate that transportation expenditure is primarily a function of socioeconomic factors while built environment measures had, on average, insignificant and at best modest influences. Breaking down the analyses into subgroups based on district urbanization characteristics, however, reveal that the influence of the built environment on travel behavior is likely to be dependent upon the regional contexts. For planners and policymakers, this implies the need to comprehensively understand how travel behavior and built environment interact in a given region or district prior to implementing neighborhood-level, built environment-focused policies aimed to influence travel behavior.

Given the overall results, it might be inferred that external factors, in the form of built environment changes, are unlikely to exert substantial effects on travel behavior for urban non-mover households. Holding the socioeconomic factor constant, this study lends a hand to the notion of stability in travel behavior, particularly and especially for urban non-movers, even in rapidly urbanizing Indonesia, where the built environment has changed considerably. A comparative analysis associating these findings against the case from other peer countries is indeed needed; however, it is somewhat unfortunate that there are no comparable studies, to the extent of literature review, that examines specifically the dynamics of household transportation spending for non-movers over a considerably long period of time (i.e., seven years) – a period where built environment changes are likely more observable than shorter period (e.g., one year). As expected, the limited number of studies on this subject is likely attributed to the lack of large scale, nationally representative panel survey in other countries.

Several policy lessons can be drawn from the results. One of the most likely relevant policy measures is the need to comprehend the travel behavior and built environment interaction in a particular regional context. Specifically, the results that indicate the built environment effect on reducing household transportation expenditure is more tractable in urban non-metro lend a hand to support land use policy that promotes compact development in small districts (*kabupaten*). This proposition does not imply that compact development holds no relevance for larger-sized or metropolitan-scale districts. The reason being is that compact development policy is likely meritorious for a range of other policy objectives beyond the quest to reducing household transportation-related expenses.

In promoting urban development that follows the principle of building up rather than spreading out, planners and policymakers can look into several practical policy options. For one, a longstanding and established planning tool as applicable in Indonesia in the form of land use code and development guidelines, i.e., RDTR (*Rencana Detail Tata Ruang*) and RTRW (*Rencana Tata Ruang Wilayah*), could help augment the efforts. Another option is through promoting the relative attractiveness of urban living as an attempt to discourage outward growth and leapfrog development patterns. In doing so, coordinated efforts that involve a wide range of stakeholders enabled by a balanced mixture of top down and bottom-up approaches are of critical importance.

Taken together, this study contributes to the travel behavior literature on several fronts, as seen from the perspective of theoretical, methodological, and empirical contributions.

Theoretical contribution. Similar to the first analytical chapter, this study further extends the application of the theoretical framework of mobility biographies to examine the dynamic aspects of travel behavior. In doing so, this study uses a balanced panel dataset derived from a large-scale, nationally representative survey to address the theoretical underpinning of factors associated with travel behavior. Results from this study, as summarized earlier, could extend the continued debate in the literature, particularly as it relates to the efficacy, viability, and under what contexts could the built environment measures influence travel behavior.

Methodological contribution. Leveraging the longitudinal character of IFLS, this study estimates the socioeconomic and built environment factors influencing travel behavior using a series of fixed effects regression. This approach allows for controlling the *omitted variable bias* and, therefore, is considered as "weak causal inference" (Endsley, 2016; Halaby, 2004). Despite its apparent advantages, to this date, there were only a few travel behavior studies that use this estimation approach (e.g., Parady et al. (2014)). To this end, this study sheds further light on the application of a family of panel regressions that travel behavior researchers might consider using in future studies to enrich the methodological applications in the field further.

Empirical contribution. Similar to the proposition of empirical contribution as specified in the first analytical chapter on the travel behavior effects of rural-urban migration, this study further advances empirical research of travel behavior in developing countries using Indonesia as the case study.

Notwithstanding the contributions of this study, a somewhat obvious limitation of this study remains. That is, the notion that no studies, to the extent of the literature review, use equivalent research design, subject of interest (i.e., urban non-movers), and data structure emanating from a longitudinal survey. This limitation somewhat hinders the quest to contextualizing the findings and methodologies of this study in the current literature.

Nonetheless, this limitation should instead provide the cue for future studies to address. In light of the growing interests of travel behavior researchers and policymakers on panel survey, owing to its potentials to address the uncharted territory and unexplored questions in the literature, more studies along a similar line as this research would hopefully emerge in the near future.

CHAPTER 6. THE INFLUENCES OF CHILDHOOD EXPERIENCES ON WALKING BEHAVIOR

6.1 Introduction

In the two preceding chapters, the emphasis was on travel behavior outcomes at the household level. The results point to the considerably modest influences of the built environment on the share of transportation expenditure from a total income. This chapter presents the relative influence of built environment exposure and socioeconomic traits on a travel behavior outcome at the level of individuals, i.e., walking behavior. A particular focus of this chapter is to assess how childhood experiences might influence behavior in the long run. This notion aligns with the urbanization phenomenon that binds the analyses of this dissertation together; that is, as the built environment and socioeconomic are evolving associated with the rapid urbanization, at what point in time do these attributes might exert influences on travel behavior?

In addition, the inquiry toward the influences of childhood experiences on travel behavior is also motivated by two additional factors. First, despite the presence of available studies investigating the association between the built environment and walking (Gehrke & Welch, 2017; Greenwald & Boarnet, 2001), the present literature tends to overlook the likely effects of prior experiences, particularly childhood experiences. Indeed, in recent studies, travel behavior scholars have lamented the importance of incorporating childhood experiences in active travel behavior modeling (Mjahed et al., 2015; Rubin, 2011; Thigpen, 2017).

Second, unpacking the potential influences of childhood experiences on travel behavior might also conceptually help address the long-standing and highly debated residential self-selection issue in travel behavior research. Residential self-selection is the notion that a given person or household would opt to live in a neighborhood that could fulfill their travel preferences (Cao et al., 2009a; Handy et al., 2006). Considering that the likelihood and influence of a given child to choose where to live in order to satiate his or her travel needs is understandably low, indirect, or even non-existent, if the built environment exposure during childhood does influence travel behavior during adulthood, it might be inferred that built environment could play a role in shaping travel behavior.

Third, the focus of Indonesia as the case study in this research is somewhat timely. A recent study indicates how little average Indonesians walk in comparison to residents in the other 45 countries surveyed, as shown in Figure 27 (Althoff et al., 2017). Specifically, the findings suggest that Indonesians tend to walk, on average, 3,513 steps a day, which is pale in comparison to Hong Kong residents that typically logged 6,880 steps a day. In fact, the study indicates that Indonesia is the least active among the countries surveyed. To this end, a careful assessment of the factors associated with walking behavior in Indonesia, including childhood experiences, might be beneficial for policy-making in the long run.

Following the introduction, this chapter is structured as follows. The next section describes a survey of the literature with an emphasis on studies that attempt to disentangle the relative importance of childhood experiences on travel behavior. The subsequent sections further elaborate on the conceptual framework and hypotheses that guide this study. A description of the data and sample characteristics, as well as the methodology, is explained consequently. Finally, a discussion of the results and its potential planning implications would conclude this chapter.

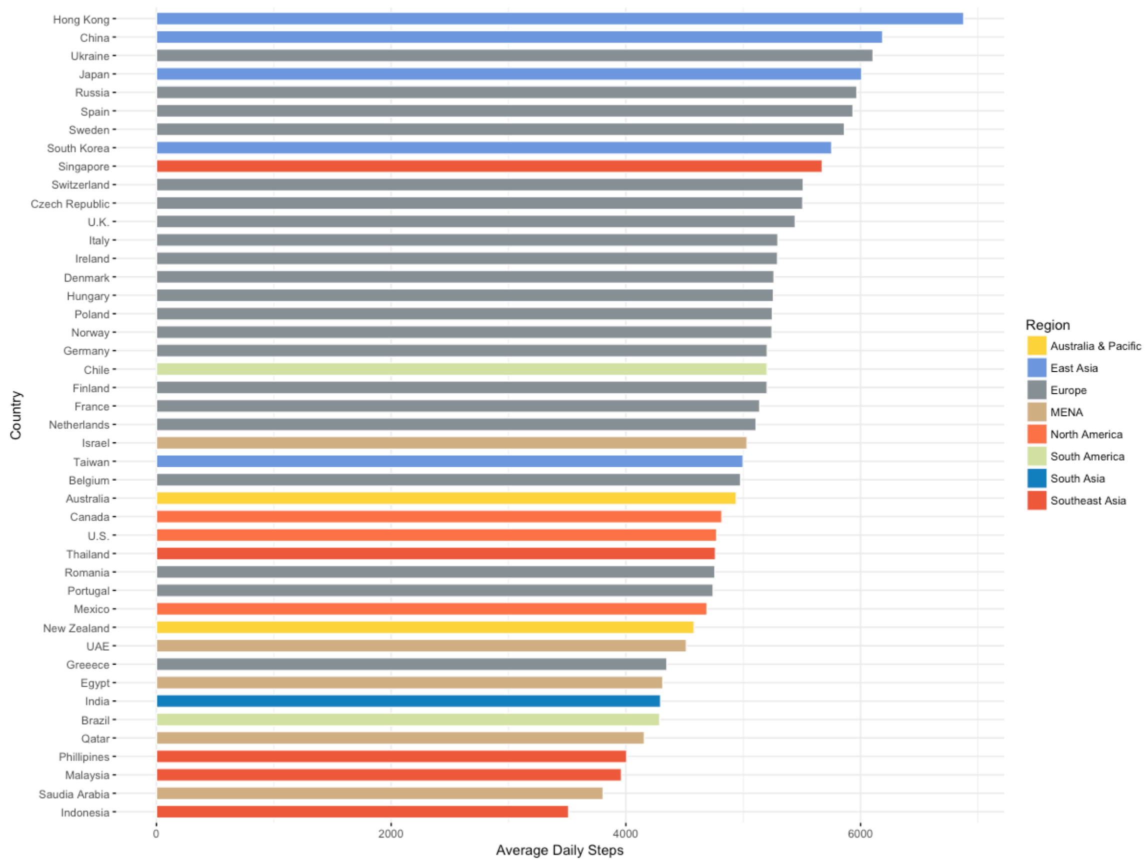

Figure 27 – Average daily steps by country. Indonesia is ranked at the bottom. (Source: Althoff et al. (2017), modified by author)

6.2 Literature Review: Childhood Experiences

Studies on the effects of childhood experiences on a variety of outcomes during adulthood using decades of panel data have long been at the epicenter in several disciplines, most visibly in public health and social sciences literature. From the public health literature, a study by Kuh & Cooper (1992) used longitudinal over the course of 36 years from a stratified sample of approximately 3,500 respondents in England, Scotland, and Wales to study the influence of childhood attributes on participation in sports and recreational activities during adulthood. Their findings suggest that certain childhood attributes, i.e., more exceptional ability to play games and energy level at 13 years, appear to influence physical activity participation during adulthood positively.

From the social sciences literature, somewhat recent studies on intergenerational mobility in the U.S. has further reinvigorated academic and policy-making interests on the value of decades of panel data to shed light on the long-term effects of childhood experiences (Chetty et al., 2014; Chetty & Hendren, 2018). Using a restricted dataset from deidentified U.S. tax records for 7 million Americans, their findings suggest the notion of 'neighborhood effects,' which is originated from childhood residential locations, as the determinants of earnings, college attendance, and fertility and marriage patterns during adulthood (Chetty & Hendren, 2018).

In the field of transportation in general and travel behavior in particular, the association between childhood experiences and travel outcomes during adulthood remains largely understudied. While transportation researchers have speculated and highlighted the importance of examining this subject (Mjahed et al., 2015; Rubin, 2011; Thigpen, 2017), the lack of transportation-related longitudinal data that examines the subjects for over a decade or more has largely inhibited further exploration. To this end, the available existing studies typically did not explicitly consider the importance of childhood experiences but instead focused on past experiences when the subjects had already entered adulthood. For example, Weinberger & Goetzke (2010) studied the effects of transit exposure from previous residences on present auto ownership using U.S. Census microdata that included the retrospective question of residences. They found that respondents who moved from US major cities served by a considerably abundant transit supply have a lower likelihood of owning a car than those who moved from small metropolitan areas. They suggested the notion of "*learned preferences*" to explain their findings.

Along a somewhat similar line, a 2015 study exploits credit card information dataset derived from a credit reporting firm for the 13-county Atlanta, Georgia region to explore the effects of past and present neighborhood characteristics on auto ownership (Macfarlane et al., 2015). The dataset provides ZIP codes information of where the respondents reside up to 9 historical locations for 227,830 completed respondents' records. Using the ZIP code-level as the geographical unit, they develop a set of built environment measures comprised of housing unit density and the non-vehicle mode share derived from the U.S. census data to estimate the past and present residential locations on auto ownership. The findings from Macfarlane et al. (2015) indicate that both past and present neighborhood characteristics play a role in influencing auto ownership. However, results from their analyses suggest that the past and present neighborhood effects on auto ownership appear to be mostly modest.

One of the few studies using decades of panel data is a recent study by Smart & Klein (2018b). Using Panel Study of Income Dynamics (PSID), they develop models to test whether past exposure to transit affects spending on transit and auto ownership in the latter years. Their findings indicate that both past and current exposure to high-quality transit environments appear to increase the likelihood of developing 'transit habit' and reducing auto ownership.

Mjahed et al. (2015) studied the role of childhood travel experience on walking behavior based on a cross-sectional online survey of 207 international respondents. The survey contains retrospective questions related to the built environment quality and parental attitudes toward walking and travel during the respondents' pre-high school years. Mjahed et al. (2015) found that respondents who grew up in pro-walk households tend to develop a stronger walking habit and preference to reside in a walkable neighborhood during adulthood. While their study has made a significant contribution to shed light on the effects of childhood experiences on travel behavior during adulthood, they appear to rely on respondents' subjective assessments. Moreover, the authors did not incorporate fine-grained built environment measures, which is likely because the survey was conducted through an online platform and the identifiable respondents' location was not recorded.

Similar to the approach as employed in Mjahed et al. (2015) study, a review of the small literature on childhood influences on travel behavior suggests that existing studies aiming to unpack the influence of childhood experiences on active transportation during adulthood use cross-sectional survey comprised of retrospective questions about individual childhood experiences (Johansson, 2005; Mjahed et al., 2015; Underwood et al., 2014; Underwood & Handy, 2012). Studies that use panel data are noticeably lacking.

6.3 Conceptual Framework and Hypotheses

6.3.1 Conceptual Framework

As discussed in the literature review section, virtually every study on the determinants of walking behavior estimates the varying factors influencing walking habits from a cross-sectional perspective. This perspective likely masks the potential influences of past experiences that could shape habit in the long run (Swait et al., 2004). As the authors suggest, there are at least three aspects that could influence behavior: '*state dependence*', or the notion that existing behavioral outcomes are influenced by previous choices; '*habit persistence*', where current behavior is affected by previous preferences; and '*initial conditions*', which refer to familiarity with a certain knowledge on particular (Swait et al., 2004). Weinberger & Goetzke (2010) summarize these aspects under the umbrella of '*learned preferences*'.

Indeed, the aspects mentioned above are related to the established Theory of Planned Behavior (TPB), which aims to comprehend the connection between cognition and behavior (Ajzen, 1991). The TPB elaborates three aspects as a predictor of behavior: *attitudes*, *subjective norms*, and *perceived behavioral control*.

Figure 28 shows the conceptual framework where childhood experiences are incorporated as a means of exploring the factors associated with walking behavior during adulthood. The orange-colored background in Figure 28 shows the dominant form of empirical analyses where existing studies tend to incorporate observable indicators during adulthood from a cross-sectional perspective solely. The inclusion of childhood experiences would be the primary contribution of this study.

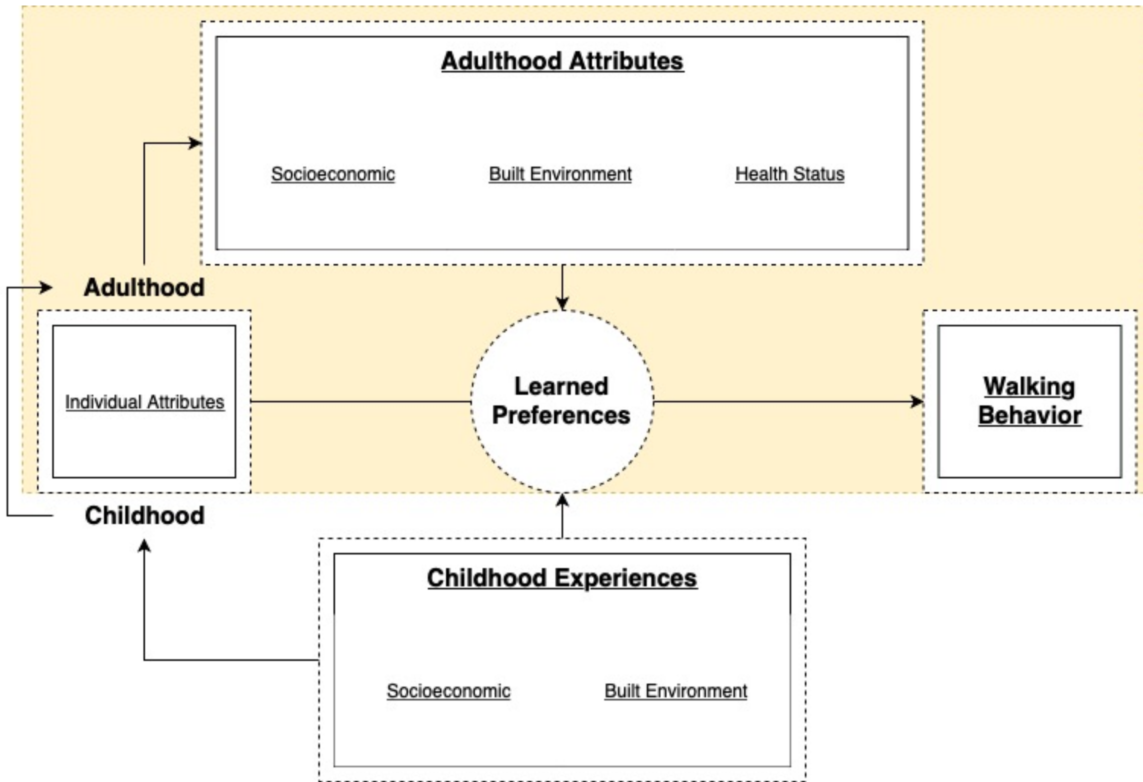

Figure 28 – Walking behavior as a function of adulthood attributes and childhood experiences

6.3.2 Hypotheses

As indicated in Figure 28 and the text box below, the first hypothesis (H_1) of this study expects that childhood experiences would significantly influence walking behavior during adulthood. In contrast, the second hypothesis (H_2) outlines that childhood experiences would not significantly exert influences on shaping walking behavior and that covariates, as measured during adulthood, would be a better predictor.

Q_3: Do socioeconomic traits and built environment exposure during childhood that form childhood experiences influence walking behavior during adulthood?

$\boldsymbol{H_{Main}}$: Childhood experiences form '*learned preferences*' that influence walking behavior during adulthood.

$H_{Alternative}$: Childhood experiences do not exert influences on walking behavior during adulthood.

To address the hypotheses as presented above, two primary factors of childhood experiences are assigned to the main variables of interest: socioeconomic and built environment exposure a given individual was exposed during childhood.

6.4 Data and Methodology

6.4.1 Data

To assemble the data, this study leverages two waves of longitudinal data provided by IFLS 3 (2000) and IFLS 5 (2014). IFLS 3 captures observable indicators of childhood experiences as recorded in 2000. IFLS 5 provides a host of information as captured in 2014 of the same individuals interviewed in 2000, or approximately 14 years apart. Similar to the sample identification approach as used in the chapter of panel analyses of urban non-movers, the analyses presented in this chapter focus on a subset of individuals who were urban over the course of their lives as recorded in IFLS 3 and IFLS 5. Using these two survey waves, socioeconomic data of the then minor individuals that were collected in 2000 are linked with the data of the same individuals that had already entered adulthood (>15 years old) in 2014. For the built environment indicators, geographical identification recorded in the IFLS at the sub-districts level (or *kecamatan*) for each individual is linked to the Village Census (PODES) data of the nearest corresponding years.

Dependent variable. The walking habit during adulthood indicator as the dependent variable is a binary indicator derived from the following question in IFLS 5: "*During the* *last 7 days, did you do any [walk] for at least 10 min. continuously?*" (module 'b3b_kk2'). The respondents who answered 'Yes' were asked to specify further whether they walked for less or more than 2 hours based on this question: "*How much time did you usually spend doing [walking] on one of those days?*" (module 'b3b_kk2'). It might be unrealistic to expect an average Indonesian to walk continuously for more than 2 hours in a single day; therefore, the working dataset consists of the subset of the overall sample that only comprised of individuals that answered either 'No' or 'Yes, for less than 2 hours' to the questions as mentioned earlier.

Independent variables. As highlighted in the conceptual framework section, the primary independent variables of interest in this chapter are built environment exposure and socioeconomic traits during childhood. The built environment factor was mainly derived from Village Census (PODES) data and included a range of physical characteristics of the neighborhoods the individuals were exposed during childhood and adulthood. A particular focus on the element of the built environment in this chapter is household density and retail density. A survey of the literature indicates that these two indicators serve as the most common proxy to capture the variation of built environment characteristics at the neighborhood level (Ewing & Cervero, 2010; Stevens, 2017), which is mostly due to the ease to obtain such metrics from administrative data. Regarding socioeconomic traits during childhood, a particular emphasis is to assess the relative importance of vehicle availability, which is an indicator that has drawn the attention of transportation researchers for decades (Anowar et al., 2014; Jong et al., 2004).

Control variables. In addition to the said primary independent variables that are assumed could influence walking behavior, various additional factors likely play a role as well. These additional factors are considered as control variables that include a range of additional socioeconomic attributes and health indicators (Table 19). As indicated, these control variables represent a socioeconomic factor that includes activity/employment status during adulthood as well as household size, presence of a minor, and income both during childhood and adulthood, as recorded in IFLS 3 and IFLS 5. Particular attention was given to the income variable as extreme outliers were observed when constructing the dataset from IFLS data. To address this issue, extreme outliers were excluded by specifying that the sample would only consist of respondents whose income was within the 1 to 99 percentage quantiles.

The health factor includes body mass index (BMI) and an assessment of self-reported health. On the one hand, the BMI indicator is provided in both survey waves. On the other hand, the self-reported health indicator is only available in IFLS 5 and represents an ordinal indicator where the respondents were asked to rank their health based on this following four choices, i.e., 1: Very healthy; 2: Somewhat healthy; 3: Somewhat unhealthy; and 4: Very unhealthy. An additional control variable includes respondents' assessment of self-reported perceived neighborhood safety. Similar to the self-reported health indicator, this variable is available only in IFLS 5. This safety indicator is an ordinal variable comprised of 4 rank-choices: 1: Very safe; 2: Safe; 3: Unsafe; and 4: Very unsafe.

The array of variables, as discussed above and listed in Table 19, highlights the benefits – and thus importance – of "true" panel data to capture individual attributes over time. In the absence of "true" panel data, attempts to probe individual attributes during his/her childhood accurately could likely be difficult. This concern is particularly relevant for this study, considering the almost 15 years gap between childhood and adulthood indicators. For instance, a regular individual might not be able to accurately recall his/her weight and height, household income or how much did his/her parents make, and household vehicle ownership during his or her childhood.

Table 19 – Walking behavior: List of variables

Variable	Description	Source	
		IFLS 3 **Childhood**	IFLS 5 **Adulthood**
Dependent variable			
Walking habit[1]	Binary indicator of a given individual's walking habit		b3b_kk2
Covariates (Socioeconomic)			
Activity/employment status[1]	Activity/employment status: 1: Employed; 2: Unemployed; 3: Student; 4: Housekeeper		bk_ar1
Income	Household monthly income (Indonesian Rupiahs - IDR)	bk_ar1; b3a_tk2	bk_ar1; b3a_tk2; b3a_re
Age (Years)	Respondent's age in 2014		bk_ar1
Household (HH) size	Total number of household members	bk_ar1	bk_ar1
HH members (< 5 years old)	Number of household members aged less than 5 years old	bk_ar1	bk_ar1
Vehicle ownership	Number of household members who own vehicle	b2_hr1	b2_hr1
Migration	Whether the respondent migrated to new residential location between 2000-2014	hhtrack	
Covariates (Health)			
Body Mass Index (BMI)	Weight (kg) / (Height (m)^2)	bus2_1	bus_us
Self-reported health	Self-reported health: 1: Very healthy; 2: Somewhat healthy; 3: Somewhat unhealthy; 4: Very unhealthy		b3b_kk1
Covariates (Built environment)			
Perceived neighborhood safety	Self-reported perceived safety:		b3a_tr
Population density	Number of populations per square kilometer	Village Census, 2008 & 2011	
Household density	Number of households per square kilometer	Village Census, 2008 & 2011	
Retail density	Number of retail establishments per square kilometer	Village Census, 2008 & 2014	
Retail-HH Balance	Number of retail establishments divided by number of households	Village Census, 2008 & 2011	

Note: [1] These variables are only measured during adulthood (> 15 years old)

In sum, following the selection of the variables and principal identification approach as outlined above, as well as considering the aim to construct a balanced dataset where the variables of interest should be available in both survey waves for each respondent, the working dataset of this study is comprised of 1073 unique individuals scattered across 14 provinces.

6.4.2 Sample Characteristics

Having constructed the working dataset, this section describes the sample characteristics by observing the descriptive statistics of the binary and continuous variables within the dataset, as shown in Table 20. For the dependent variable of walking habit, approximately 60% of the respondents mentioned they maintained walking habits in the past week at the time of the survey. This value might be considered as welcome news given how little an average Indonesian walk based on a recent study (Althoff et al., 2017).

In terms of socioeconomic indicators, several observations of income, household size and composition, vehicle ownership, and individual traits point to the dynamics of socioeconomic indicators separated over 14 years. For instance, Table 20 shows the relative average increases of respondents' wealth where inflation-adjusted average income increased by approximately IDR 3 million (~U$D 220, August 2019 exchange rate) and average vehicle ownership increased by 0.43 between 2000 and 2014.

The notion of increasing wealth for average Indonesians, as indicated in income and vehicle ownership variables, is tractable as well from the indicator of Body Mass Index (BMI). As shown in Table 20, the average BMI of Indonesian children was 15.48, which is categorically underweight. Fourteen years ahead following the same individuals, the average BMI of Indonesian adults was 21.81, which is considered as an average weight.

Table 20 – Walking behavior: Descriptive statistics of the sampled individuals, 2000-2014

Variable		Descriptive Statistics			
		Mean	S.D.	Min	Max
Dependent variable					
Walking habit	C				
	A	0.61	0.49	0	1
Socioeconomic					
Income (IDR)[1]	C	1,247,739.53	1,028,557.56	81,503.98	7,141,002.64
	A	5,168,175.79	4,188,668.13	440,000.00	32,000,000.00
Age (Years)	C				
	A	20.64	4.03	15	30
Household (HH) size	C	5.64	2.16	1	15
	A	5.00	2.00	1	13
HH members (< 5 years old)	C	0.95	0.80	0	4
	A	0.39	0.65	0	3
Vehicle ownership	C	1.01	1.11	0	9
	A	1.44	0.99	0	6
Household head	C				
	A	0.03	0.18	0	1
Gender	C				
	A	0.51	0.50	0	1
Migrant	C				
	A	0.02	0.15	0	1
Health					
Body Mass Index (BMI)	C	15.65	2.58	7.11	44.64
	A	21.90	4.74	13.70	45.77
Built environment					
Population density	C	7,283.43	8,295.19	1,010.50	48,212.47
	A	7,992.54	7,096.58	1,072.04	48,940.08
Household density	C	1,657.10	1,959.30	225.60	11,665.90
	A	2,097.29	1,913.48	295.60	11,698.80
Retail density[2]	C	2.83	2.11	0.32	12.95
	A	5.25	7.74	0.03	48.98
Retail-HH Balance	C	0.0031	0.0027	0.00030	0.0161
	A	0.0020	0.0020	0.00009	0.0156

Note:
C: Childhood; A: Adulthood.
[1] Income in 2000 is adjusted to reflect its value in 2014 using the Consumer Price Index (CPI)
[2] Street vendors are not documented since Village Census 2000 does not provide such information.

One health indicator not included in Table 20 is self-reported health due to its non-continuous nature. From 1073 surveyed individuals in the sample, 233 assessed they were ‘Very healthy’, 672 answered ‘Somewhat healthy’, 161 considered they were ‘Somewhat unhealthy’, and only 7 respondents believed they were ‘Very unhealthy’. Owing to the considerably small number of respondents who answered their health condition was ‘Very unhealthy’, this observation is collapsed to the sample who answered they were ‘Somewhat unhealthy’.

Regarding individual traits, the data has a reasonably proportionate share between males and females. Only three percent of the sample were household head, which is understandable considering the average age of the sample was 20.64 years – well below the typical age at first marriage, i.e., for women it was 23.1 and for men it was 27.5, as captured in the 2015 Intercensal Population Survey (Qibthiyyah & Utomo, 2016). Moreover, approximately two percent of the individuals surveyed had relocated to different residential places throughout their lives between 2000 and 2014.

In terms of the built environment, in general, the data suggests that Indonesia’s urban environment is getting more populated. The average population and household density increased by around 709.11 people/km^2 and 440.19 household/km^2, respectively. Furthermore, as urban areas getting denser, retail opportunities tend to proliferate as the average retail density increased by 2.43 facilities/km^2 between 2000 and 2014.

A particular built environment indicator that is not listed in the descriptive statistics (Table 20) is perceived neighborhood safety since this indicator is not a continuous variable. From the sample of 1073 individuals, 208 considered the neighborhood he/she lives in was ‘Very safe’, 770 perceived it was ‘Safe’, 88 thought it was ‘Unsafe’, and only 7 deemed it was ‘Very unsafe’. Since the number of respondents who answered that the neighborhood was ‘Very unsafe’ was relatively small in comparison to the entire sample, this observation is collapsed to the ones answered ‘Unsafe’.

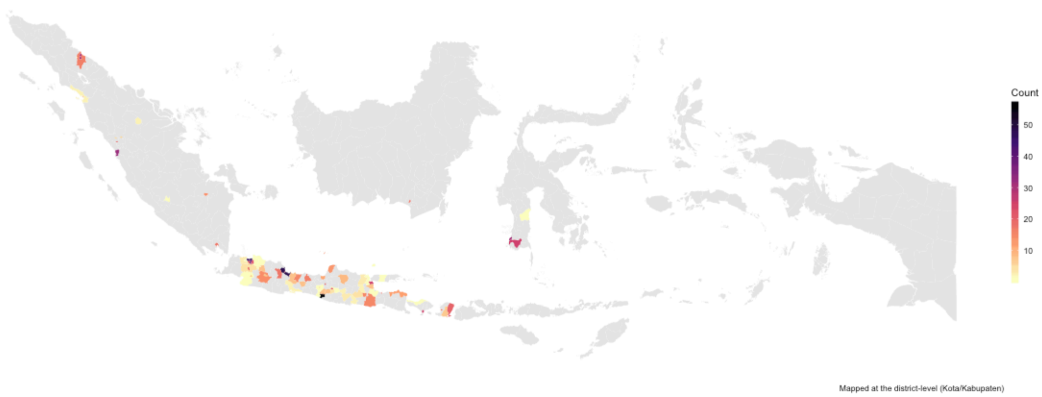

Figure 29 – Geographic distribution of the respondents based on their residential location in 2014 (District-level)

Table 21 – Provincial distribution of the respondents based on their residential location in 2000 and 2014

Province	Region	2000		2014	
		# Indiv.	%	# Indiv.	%
Bali	Bali & NT	26	2.42	26	2.42
West Nusa Tenggara	Bali & NT	43	4.01	43	4.01
Banten	Java		0.00	7	0.65
Central Java	Java	160	14.91	160	14.91
East Java	Java	116	10.81	118	11.00
Jakarta	Java	205	19.11	199	18.55
West Java	Java	174	16.22	176	16.40
Yogyakarta	Java	81	7.55	80	7.46
South Kalimantan	Kalimantan	17	1.58	17	1.58
South Sulawesi	Sulawesi	57	5.31	57	5.31
Lampung	Sumatera	18	1.68	17	1.58
North Sumatera	Sumatera	96	8.95	95	8.85
Riau	Sumatera	2	0.19	2	0.19
South Sumatera	Sumatera	16	1.49	15	1.40
West Sumatera	Sumatera	62	5.78	61	5.68
	Total	1073	100	1073	100

Figure 29 and Table 21 illustrate the geographic distribution of the sample at the districts (*kabupaten*) and province-level, respectively. As indicated, most of the respondents resided in Java island, particularly in the country's densely populated regions such as Jakarta and Yogyakarta. Outside of Java, a fair share of the surveyed individuals calls the city of Medan in North Sumatera, Padang in West Sumatera, and Makassar in South Sulawesi as their home.

6.4.3 Estimation Approaches

The estimation approaches to model the factors associated with walking behavior are determined by the binary nature of the dependent variable, i.e., whether a given individual stated they walked continuously for at least 10 minutes but less than 2 hours in the past week or not. To this end, a binary logistic regression estimator is used following Equation 11 below (Ben-Akiva & Lerman, 1985; Rumbach & Shirgaokar, 2017):

$$P_n(i) = \frac{e^{\mu V_{in}}}{e^{\mu V_{in}} + e^{\mu V_{jn}}} \tag{11}$$

The specification of binary logistic regression includes P_n = probability of walking, n = individual indicator, i = walking, j = did not walk, μ = utility, V_{in} = measured variables associated with walking i. In this study, the measured variables include the built environment, socioeconomic, and health indicators.

In addition to logistic regression, considering that the respondents within the sample are clustered in particular geographies/administrative boundaries, mixed-effects logistic regression is explored. This method is recommended as a viable approach to account for a potential clustering effect or when both fixed and random fixed effects are present (Boisjoly et al., 2017; Nasri & Zhang, 2012). With regard to this case, it is reasonable to expect that a given person might exhibit similar walking behavior with an otherwise similar individual living in the same places or neighborhoods.

A log-likelihood ratio test (Equation 12 (Garrow, 2013)) is applied to evaluate the performance of each estimation approach where LL_R is the log-likelihood value of logistic regression model and LL_U is the log-likelihood value of mixed-effects logistic regression model. Consequently, and following the principle of parsimony, the decision regarding which regression approach to use would follow the log-likelihood ratio test.

$$-2[LL_R - LL_U] \sim \chi^2_{NR,\alpha} \quad (12)$$

Given the results as presented in the log-likelihood ratio test, the chosen regression approach, i.e., either logistic regression or mixed-effects logistic regression, is used to estimate a series of models, as shown in Table 22. As indicated, these models include the childhood-only model, adulthood-only model, and the combination of both.

Table 22 – Models to be estimated

Group	Model	Description
1	1	Childhood-only model
	2	Adulthood-only model
2	3	Childhood + Adulthood (Full model)
	4	Childhood + Adulthood (Partial model)

The primary aim of developing these models is to test the hypotheses as outlined previously. In doing so, and similar to the process of evaluating which regression approach would be more appropriate, an evaluation of the model performance relative from one to another is conducted using log-likelihood ratio tests. The tests include assessing 1) whether the childhood-only model outperforms adulthood-only model and 2) whether the combination of childhood and adulthood indicators outperform adulthood-only model. The childhood-only (Equation 13) and adulthood-only (Equation 14) models are described as follow:

$$Walk(P_n) = \beta_0 + \beta_1 SE_{Child} + \beta_2 BE_{Child} + \beta_3 BMI_{Child} + \beta_4 V + e \tag{13}$$

$$Walk(P_n) = \beta_0 + \beta_1 SE_{Adult} + \beta_2 BE_{Adult} + \beta_3 BMI_{Adult} + \beta_4 V + e \tag{14}$$

where, P_n = probability of walking; SE = Socioeconomic traits that include household income, household size, number of employed household members, number of household members aged five years old or less, and vehicle ownership; BE = Built environment indicators, BMI = Body Mass Index; and V = Additional covariates measured during adulthood in 2014 and are incorporated in each model that include personal (i.e., age, gender, employment, self-reported health) and household-level characteristics (i.e., migration record).

For the model that combine childhood and adulthood indicators, I estimate this following 'full model' (Equation 15):

$$\begin{aligned} Walk(P_n) = {} & \beta_0 + \beta_1 SE_{Child} + \beta_2 BE_{Child} + \beta_3 BMI_{Child} + \beta_4 V + \beta_0 \\ & + \beta_1 SE_{Adult} + \beta_2 BE_{Adult} + \beta_3 BMI_{Adult} + \beta_4 V + e \end{aligned} \tag{15}$$

This estimation could likely suffer from multicollinearity issues, especially for the built environment indicators, where a given neighborhood physical characteristics might be relatively similar across the years. To address this potential problem, this study explores the ‘partial model’ where the built environment covariates during adulthood are excluded.

6.5 Results and Discussion

6.5.1 Model Comparison

This section presents the analysis of the log-likelihood ratio test from logistic regression and mixed-effects logistic regression[7] for the childhood-only model as an approach to evaluate which regression techniques would be more appropriate (in the Appendix shows the regression result). Table 23 presents the test results, and it is apparent that the log-likelihood ratio test is statistically significant (Prob(>chisq) = 0.07636) and therefore indicates that the mixed-effects model is a more appropriate option, which also suggests that the observations are clustered[8]. To this end, in conducting the analyses for each of the models listed in Table 22, mixed-effects logistic regression would be the model of choice.

Table 23 – Log-likelihood ratio test: regular logistic regression and mixed-effects logit

Model[1]	# Attributes	Log-likelihood	chisq	Prob(>chisq)
Logistic regression	19	-702.33		
Mixed-effects logit	20	-700.76	3.1407	0.07636*
Note:				*p<0.1; **p<0.05; ***p<0.01

[7] The mixed-effects logistic regressions were estimated in R Studio platform using ‘lme4’ package (Bates et al., 2019).

[8] A series of mixed-effects logistic regression models was developed to assess at which geographic level should the model is clustered. Results from estimating a childhood-only model indicate clustering the estimation at the province level (LL=-699.97) offers better goodness of fit than when the estimation was clustered at the district (LL=-701.28) or sub-district level (-701.18).

6.5.2 Results

This section presents results from a series of mixed-effects logistic regressions following the list of the models to be estimated as outlined in Table 22. First and foremost, Table 24 describes estimation results from the childhood- and adulthood-only model, and Table 25 presents an analysis of which model has a better goodness-of-fit. It is apparent from Table 25 that the childhood-only model offers a better model fit than the adulthood-only model where a statistically significant log-likelihood ratio test is observed (Prob(>chisq) = 0.0000). The smaller Akaike information criterion (AIC) and Bayesian information criterion (BIC) values of the childhood-only model in comparison to the adulthood-only model further corroborate the proposition that the former has a better goodness-of-fit than the later.

Comparing the coefficients of each model (Table 24) sheds further light on the factors influencing walking behavior. Given the particular focus on the potential influences of childhood experiences, the results indicate the relative importance of socioeconomic and built environment measures during childhood in comparison to experiences during adulthood. For instance, it appears that household density during childhood exerts a statistically significant influence on walking behavior during adulthood, while the same measure during adulthood does not. Specifically, the estimated coefficient suggests that the likelihood to maintain a walking habit during adulthood increase by 0.01 percent for each additional increase in household density, all else equal.

Table 24 – Model results: Walking behavior, childhood- and adulthood-only model

	Childhood-only (1)		Adulthood-only (2)	
	Estimate	Std. Error	Estimate	Std. Error
Childhood				
Household Density	0.0001*	(0.00005)		
Retail Density	-0.028	(0.038)		
Vehicle availability	-0.098*	(0.060)		
Income (logged)	-0.086	(0.089)		
Body Mass Index (logged)	0.341	(0.455)		
Household size	-0.017	(0.034)		
# Household members < 5 years old	0.175*	(0.102)		
Adulthood				
Household Density			0.00003	(0.0001)
Retail Density			0.005	(0.012)
Vehicle availability			0.004	(0.069)
Income (logged)			0.104	(0.094)
Body Mass Index (logged)			0.099	(0.342)
Household size			0.009	(0.044)
# Household members < 5 years old			-0.203	(0.128)
Migration (0 = Did not migrate)	-0.384	(0.429)	-0.404	(0.428)
Gender (0 = Female)	0.276**	(0.137)	0.271**	(0.138)
Household head (0 = Not a head)	0.040	(0.392)	0.218	(0.404)
Age	0.025	(0.025)	-0.0002	(0.022)
Health (Ref.: Somewhat unhealthy)				
Very healthy	-0.044	(0.217)	-0.043	(0.215)
Somewhat healthy	0.216	(0.184)	0.236	(0.183)
Employment (Ref.: Unemployed)				
Employed	0.201	(0.220)	0.168	(0.220)
Housekeeper	0.514*	(0.296)	0.630**	(0.302)
Student	0.470**	(0.234)	0.403*	(0.233)
Safety (Ref.: Very safe)				
Safe	-0.111	(0.168)	-0.107	(0.168)
Unsafe	-0.483*	(0.264)	-0.464*	(0.263)
Constant	-0.224	(1.651)	-1.889	(1.651)
Observations	1,073		1,073	
Multilevel variance comp.	0. 04602		0. 05908	
Log Likelihood	-700.762		-703.539	
Akaike Inf. Crit.	1,441.524		1,447.077	
Bayesian Inf. Crit.	1,541.088		1,546.642	
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 25 – Log-likelihood ratio test: Adulthood- and childhood-only model

Model	# Attributes	Log-likelihood	chisq	Prob(>chisq)
Adulthood-only (2)	20	-703.54		
Childhood-only (1)	20	-700.76	5.3537	0.0000***
Note:				*p<0.1; **p<0.05; ***p<0.01

A measure of vehicle availability suggests that for each additional unit increase in the number of the household member that owns a vehicle during childhood, the likelihood of developing walking habit during adulthood decrease by 9.3 percent (*p-value* < 0.1), holding other variables constant. The same indicator during adulthood exhibits an insignificant association (*p-value* > 0.1). It actually has a positive coefficient sign, which is at odds with the one observed in the childhood-only model.

The childhood-only model also predicts that the presence of a non-adult sibling might stimulate walking behavior during adulthood. Specifically, the estimated coefficients suggest that for every additional increase in the number of a household member aged five years or less during childhood could increase the likelihood of maintaining walking habit during adulthood by approximately 19.1 percent, all else equal. This might imply that a range of activities enabled by the presence of peer individuals at similar age during childhood might influenced walking behavior later in life.

Several control variables that represent a combination of time-invariant and present indicators indicate largely expected findings. For instance, in terms of employment status, both childhood- and adulthood-only models predict that being unemployed reduces the likelihood of a given adult to maintain a walking habit. Specifically, it appears that the respondents who worked as a housekeeper are 67.2 (*p-value* < 0.1) to 87.7 (*p-value* < 0.05) percent more likely to maintain a walking habit than an otherwise similar but unemployed individual, holding other variables constant. A similar observation is observed for the respondents who were a student where their likelihood to sustain a walking habit is 49.6 (*p-value* < 0.1) to 59.9 (*p-value* < 0.05) percent more likely than unemployed adults, all else equal.

Table 24 also reports the relative importance of neighborhood safety in influencing walking behavior, as indicated in both childhood- and adulthood-only models. The coefficient from both models suggests that the likelihood of a given adult to maintain a walking habit decrease by approximately 59 to 62 percent if the individual felt the neighborhood was unsafe in comparison to those who felt the neighborhood was very safe, holding other variables constant.

Having estimated and discussed the results from childhood- and adulthood-only model (Table 24), Table 26 presents the estimation results where both childhood and adulthood indicators are estimated simultaneously, forming the full model. Estimation results from the full model indicate that several childhood indicators are no longer maintain statistically significant influences. This is most evident in terms of household density and the number of household members aged five years old or less, as shown in Table 26 (Model 3). Only one variable maintains its significance (*p-value* < 0.1) and a consistent coefficient sign, i.e., vehicle availability, across the childhood-only, adulthood-only, and full model. Due to this reason, the model that excludes several adulthood indicators, i.e., the partial model is estimated, as shown in Table 26 (Model 4).

Table 26 – Model results: Walking behavior, full and partial model

	Full model (3)		Partial model (4)	
	Estimate	Std. Error	Estimate	Std. Error
Childhood				
Household density	0.0001	(0.0001)	0.0001*	(0.00005)
Retail density	-0.025	(0.039)	-0.025	(0.038)
Vehicle availability	-0.111*	(0.061)	-0.104*	(0.060)
Income (logged)	-0.143	(0.094)	-0.126	(0.092)
Body Mass Index (logged)	0.315	(0.497)	0.276	(0.495)
Household size	-0.004	(0.036)	-0.018	(0.034)
# Household members <5 years old	0.168	(0.105)	0.175*	(0.102)
Adulthood				
Household density	-0.00003	(0.0001)		
Retail density	0.003	(0.012)		
Vehicle availability	0.023	(0.070)	0.016	(0.069)
Income (logged)	0.172*	(0.099)	0.142	(0.093)
Body Mass Index (logged)	0.062	(0.377)	0.109	(0.373)
Household size	-0.020	(0.048)		
# Household members <5 years old	-0.171	(0.132)		
Migration (0 = Did not migrate)	-0.451	(0.437)	-0.402	(0.430)
Gender (0 = Female)	0.279**	(0.139)	0.285**	(0.137)
Household head (0 = Not a head)	0.147	(0.415)	0.156	(0.399)
Age	0.023	(0.025)	0.020	(0.025)
Health (Ref.: Somewhat unhealthy)				
Very healthy	-0.049	(0.218)	-0.057	(0.217)
Somewhat healthy	0.230	(0.186)	0.221	(0.185)
Employment (Ref.: Unemployed)				
Employed	0.177	(0.223)	0.183	(0.221)
Housekeeper	0.612**	(0.303)	0.527*	(0.298)
Student	0.452*	(0.237)	0.462**	(0.235)
Safety (Ref.: Very safe)				
Safe	-0.153	(0.170)	-0.121	(0.168)
Unsafe	-0.497*	(0.267)	-0.494*	(0.264)
Constant	-1.997	(2.010)	-1.907	(1.989)
Observations	1,073		1,073	
Multilevel variance comp.	0.05791		0.04314	
Log Likelihood	-697.339		-699.331	
Akaike Inf. Crit.	1,448.679		1,444.662	
Bayesian Inf. Crit.	1,583.091		1,559.161	
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 27 – Log-likelihood ratio test: Full and partial model

Model	# Attributes	Log-likelihood	chisq	Prob(>chisq)
Full model (3)	27	-697.339		
Partial model (4)	23	-699.331	3.983	0.4083
Note:				*p<0.1; **p<0.05; ***p<0.01

Using the estimation approach of the partial model (Model 4, Table 26), it appears that several childhood indicators (i.e., household density, vehicle availability, and the number of household members aged five years old or less) reclaim the significance and a more or less similar estimated coefficients as indicated in the childhood-only model (Model 2, Table 24).

Finally, owing to the overall objective of this chapter was to assess the relative importance of childhood experiences, a log-likelihood ratio test is estimated to assess whether the full model (Model 3, Table 26) that incorporates both childhood and adulthood indicators perform relative to the adulthood-only model (Model 2, Table 24). The results from the log-likelihood ratio test, as shown in Table 28, indicate that adding childhood indicators improve the predictive strength of the model (Prob(>chisq) = 0.08819) and therefore substantiate the advantages of incorporating childhood experiences to estimate the factors associated with walking behavior during adulthood.

Table 28 – Log-likelihood ratio test: Adulthood-only and full model

Model	# Attributes	Log-likelihood	chisq	Prob(>chisq)
Adulthood-only (2)	20	-703.54		
Full model (3)	27	-697.339	12.399	0.08819*
Note:				*p<0.1; **p<0.05; ***p<0.01

6.6 Conclusion

This chapter addresses the third and last research questions of this dissertation on the potential influences of childhood experiences on travel behavior during adulthood, specifically walking habit. By assembling a 14-year worth of data capturing the life-course of 1073 urban individuals from their childhood (2000) to adulthood (2014), this study provides suggestive evidence on the influences of childhood experiences on walking habits during adulthood.

Specifically, this study finds that exposure to dense urban environments during childhood could induce a given individual to maintain a walking habit during adulthood. One might imagine that growing up in a dense urban environment might afford the development of tightly knit social networks (e.g., friendship with peers) where children could access without having to rely on other adults to help them transport themselves accessing it.

In addition, this study also discovers the association between vehicle availability and walking behavior where greater access to a vehicle during childhood could reduce the propensity of a certain individual to develop walking habits during adulthood. In sum, these findings lend a hand to support the notion of 'learned preferences' as specified in the hypothesis (Figure 28).

Considering these findings, several planning implications can be drawn. For one, the results highlight the relative importance of the development of a compact and connected urban environment as it could wield a long-term influence on travel behavior. Policies aimed to encourage such development patterns might need to be put in places and or further promoted. For instance, policymakers could put a greater emphasis on supporting the principle of upward growth rather than leapfrog development pattern. This can be achieved through several urban design support systems and planning tools, including the existing ones as applicable in Indonesia, e.g., land use code and urban development guidelines (RDTR (*Rencana Detail Tata Ruang*) and RTRW (*Rencana Tata Rung Wilayah*)). Along a similar line, policymakers could also promote the relative attractiveness of urban living intended particularly for parents. The aim is to nudge residential locational decisions toward increasing the propensity of parents to opt to live in a dense and spatially mixed urban environment instead of peripheral and inaccessible suburban locations.

The second planning implication hinges on the apparent influence of vehicle availability on walking behavior. As exposure to a higher number of vehicle availability during childhood could hinder the development of walking habits during adulthood, policies aimed to reduce personal vehicle dependency might need to be pursued as it could wield a lasting influence in shaping travel behavior. In doing so, planners and policymakers could consider and adopt several planning and policy tools that revolve around lowering both the incentives and needs for owning private vehicles. These include improving access to modern and safe transit systems, ameliorating the built environment that could reduce the propensity and needs of owning automobile or two-wheelers, transportation demand management, among others. Understandably, achieving that objective is by no means an easy task considering the overall climate in favour of vehicle ownership in Indonesia, shaped by years of political economy decisions that encourage the ownership and use of motorized transport (Gaduh et al., 2017; Hook & Replogle, 1996). Nonetheless, in light of the findings in a recent study that posits how average Indonesians walk the least relative to other countries as well as considering the potential negative implications that could follow due to sedentary lifestyle, the needs to take a concrete action to encourage increased walking habit are of significant importance.

Taken together, this study makes three primary contributions to the travel behavior literature, as seen from the perspective of theoretical, methodological, and empirical contributions.

Theoretical contribution. Following the notion of "*learned preferences*" influencing present behavior, this study contributes to augmenting the theoretical proposition on the origin of travel behavior. Concurrently, it also further complicates the debate on the extent of residential self-selection in the travel behavior literature. In regard to that, this study paves the path for future studies on the substantive topic of the influences of past experiences on present travel behavior. More studies are therefore needed to test whether the theoretical proposition on the relative importance of past experiences emanating from this study is applicable in different contexts.

Methodological contribution. Leveraging the longitudinal nature of the IFLS data, this study offers a methodological contribution where years-worth of panel surveys can be used to assess the relative importance of childhood experiences in shaping travel behavior in comparison to the adulthood-only model. This methodology is replicable and indeed should be extended and tested in future studies where researchers might follow the life-course of sampled individuals from their childhood to adulthood to tease out the influences of past experiences on present travel behavior.

Empirical contribution. Along the line of empirical contribution as specified in the first and second analytical chapters of this dissertation, this study further advances empirical research of travel behavior in developing countries using Indonesia as the case study. The case of Indonesia is also particularly timely from a policy standpoint, considering a relatively recent study indicates that Indonesians walk the least in comparison to residents in several other countries (Althoff et al., 2017).

Despite the contributions that this study could make, several limitations persist. One of the most visible limitations is related to the fact that this study mainly concerns with the walking habit of primarily young adults., where the maximum age of the respondents in the sample was 30 years old. It remains unclear whether, and to what extent, childhood experiences might exert influences on walking behavior of older adults living in their 30s and beyond. The other limitation stems from the notion that there are no equivalent studies, to the extent of the literature review, that uses similar research design and longitudinal data structure. This condition makes it somewhat challenging to compare and contrast the findings from this study.

These limitations could instead offer an avenue for future studies to address. For one, the availability of the next wave of IFLS data with consistent information from its predecessors might allow exploration of childhood experiences on the walking behavior of older adults. Furthermore, this study could hopefully ignite researchers and policymakers' interests on the relative importance of childhood experiences on travel behavior. The increasing interest in this subject might eventually pave the way for the prevalence of multiyear panel surveys that can be used to address the inquiries along the line of the role of past experiences, particularly childhood, on travel behavior in a later stage of individuals' lives.

CHAPTER 7. CONCLUSIONS

7.1 Summary

Rapid urbanization is the prominent phenomenon of this century. Developing Asia, including Indonesia, had seen a sustained urbanization process that accelerated rapidly relative to the ones observed in Europe and North America in previous decades (Asian Development Bank, 2012; World Bank, 2018b). This rapid urbanization process entails substantial changes in the built environment and socioeconomic characteristics.

Linking the urbanization phenomenon with travel behavior literature, which considers the built environment and socioeconomic as primary determinants influencing its variation, presents a compelling case to examine the potential synergistic relationship between the changes in the built environment and socioeconomic associated with urbanization and the dynamics of travel behavior over time. In light of the considerably rapid urbanization process occurring in Indonesia, the country presents a compelling case to observe and analyze travel behavior dynamics.

Following two of the three aspects that make up the urbanization process (Figure 5), i.e., rural-urban migration, natural growth, and area reclassification (World Bank, 2018a), this dissertation explored three research questions directly linked with rural-urban migration and evolving built environment and socioeconomic characteristics associated with natural growth:

1) How is a household's travel behavior affected by a move from rural to urban location?

2) For urban non-movers, how do the dynamic of the built environment and socioeconomic changes influence travel behavior?
3) Do socioeconomic traits and built environment exposure during childhood that forms childhood experiences influence walking behavior during adulthood?

In addressing these research questions, this dissertation borrows the theoretical framework of mobility biographies (Lanzendorf, 2003; Scheiner, 2007). This theoretical framework considers the dynamic aspects of travel behavior associated with the life-course or events a given household or individual might encounter.

The preceding three analytical chapters have addressed these research questions and provided empirical estimates to shed light on the factors associated with the changes, stability, and, to some extent, the origin of travel behavior under the backdrop of rapid urbanization. Chapter 4 addresses the travel behavior effects of relocating from rural to an urban environment. It finds that relocating to urban areas could reduce household transportation expenditure, as a proxy for travel demand, by approximately 10% relative to the ones who remained rural.

Chapters 5 and 6 focus primarily on the natural growth aspect of the urbanization process. Specifically, while Chapter 4 focuses on relocating households, Chapter 5 estimates the influences of the built environment and socioeconomic changes on the travel behavior of urban non-mover households. It finds the modest, inelastic, and insignificant relationship between gross population density and household transportation expenditure, which also speaks to the relative stability of travel behavior for non-movers even when the built environment they live in evolves considerably rapidly.

Chapter 6 seeks to disentangle at which life stage could built environment influences present walking behavior. It finds that greater exposure to dense environments during childhood could induce a walking habit during adulthood.

7.2 Dissertation Contributions

Taken together, this dissertation offers theoretical, methodological, and empirical contributions to the travel behavior and built environment interaction literature, in particular, and the field of city and regional planning, in general.

7.2.1 Theoretical Contributions

From the lens of theoretical contribution, this dissertation extends the application of a life-course approach, i.e., mobility biographies, to observe the association between travel behavior and the evolving characteristics of both macroscopic (e.g., urban form) and microscopic (e.g., household socioeconomic) factors (Kitamura, 1990). To date, the prevailing theoretical framework applied in the majority of travel behavior studies tends to overlook these dynamic aspects. This dissertation offers a refreshed perspective and shows the relative importance of mobility biographies to guide the observation of the changes, stability, and, to some extent, origin of travel behavior.

Furthermore, results from the three analytical chapters of this dissertation collectively highlight the notion of 'windows of opportunity' (Müggenburg et al., 2015; Prillwitz et al., 2006), where travel behavior might be shaped through life events, specifically residential relocation in the form of rural-urban migration, and past experiences. Recognizing the notion of 'window of opportunity', future studies might adopt this proposition to guide the research framework and study design, as well as expectation or explanation of the findings.

7.2.2 Empirical Contributions

The empirical contributions of this dissertation largely stem from the case being analyzed and the primary dataset that enables the empirical estimations. That is, using Indonesia as the case study expands the present literature that has so far been dominated by empirical studies from developed and more prosperous countries. In that regard, Indonesia offers a compelling case not only because of its present status as lower middle-income country, which remains mostly understudied in travel behavior literature but also because the country's considerably rapid rural-urban transformation offers a well-suited background phenomenon to analyze travel behavior dynamics guided by the theoretical frameworks as presented earlier.

The use of a large-scale, longitudinal survey dataset, i.e., Indonesian Family Life Survey (IFLS), as the primary data source, further highlights the empirical contributions of this dissertation from the data application standpoint. As indicated in Chapter 2: Literature Review, the use of this type of data in travel behavior, in particular, and planning literature, in general, remains noticeably lacking but gradually growing in recent years. This dissertation contributes to filling this gap on the lack of longitudinal studies and, in doing so, tackles research questions that otherwise would remain unaddressed had researchers continue to rely on the conventional cross-sectional data. The opportunity to address unexplored empirical research questions, therefore, naturally opens up the potential adoption of considerably novel empirical strategy and methodological approaches that have not gained prevalence in the field of travel behavior, as indicated in the following sub-section.

7.2.3 Methodological Contributions

From the lens of methodological contributions, this dissertation shows the exploration of a series of quantitative methods that remain relatively underutilized in travel behavior and planning literature. As shown in the preceding analytical chapters, this exploration of a series of quantitative methods is enabled by the longitudinal nature of the primary data source. This notion is particularly apparent in Chapter 4, addressing the substantive topic of rural-urban migration. In this chapter, the methodological approach leverages the unique characteristic of the survey dataset that tracks the sampled households and individuals over time, even if the households or individuals in question relocated to different places. Assigning the relocating households as treatment and those who remained or relocated to rural as control, which is identified through propensity score matching, the causal inference approach of Difference-in-Differences (DID) regressions are estimated to assess the impact of the exposure to urban environments on household transportation expenditure. To the extent of review on travel behavior literature, there are no previous studies that use this empirical strategy and methodological approach combining propensity score matching and DID regressions on micro-level household panel data.

The methodological contributions of analytical Chapters 5 and 6 are also enabled by the use of longitudinal data. In Chapter 5, fixed effects panel regressions are estimated on the sample of urban non-mover households. In Chapter 6, results from mixed-effects logistic regressions are presented. The application of these estimation methods in travel behavior studies does not as markedly lack as the approach adopted in Chapter 4. Nonetheless, to date, there is only a handful of relevant travel behavior studies that use the said estimation approaches, e.g., studies using fixed effects panel regression include Parady et al. (2014); while, the ones using mixed-effects regressions include studies by Nasri and Zhang (2012) and Wasfi et al. (2017).

7.3 Policy and Research Implications

7.3.1 Policy Implications

Collectively and individually, findings from each chapter of this dissertation offer several policy lessons related to transportation and urban development. For one, the results, indicating that 'sudden' exposure to the denser environment could induce travel behavior changes, suggest the relative importance of ensuring the supply of compact and connected environment, which supports the proposition of building up, e.g., through densification, rather than spreading out or outward growth. This proposition also holds relevance in light of the aim of promoting a more sustainable travel pattern through increased walking as continued exposure to dense urban environments during childhood appears to induce greater walking habits during adulthood.

7.3.2 Research Implications

This dissertation offers potential research implications that could apply to future studies. First, and likely the most obvious, this dissertation highlights the merit of using large-scale longitudinal survey data, which span over years or even decades, to tackle research questions that otherwise might not be possible to be addressed using a conventional cross-sectional dataset. Second, if the research objective revolves around the association between residential relocation and travel behavior, then the survey methodology in question should be designed in a way that would allow tracking the sampled respondents to capture the dynamics of internal or even external migration, which by itself is a global phenomenon.

Devouring into the efforts of proliferating longitudinal approach in travel behavior literature is likely not without challenges. One of the most likely challenges is the time, labor, and monetary resources required to institute multiyear panel surveys. The second challenge relates to the attrition issue that has long been considered as the “Achilles heel” of conducting a panel survey (Thomas et al., 2001, p. 559, 2012, p. 109). Nonetheless, researchers have lamented the cost-saving potential associated with running panel surveys in comparison to conducting a repeated cross-sectional observation as well as several strategies to minimize attrition (Thomas et al., 2001, 2012; Tourangeau et al., 1997).

APPENDICES

A.1 Appendix – Chapter 3: Research Context and Framework

A.1.1 Analysis of Indonesia's Urbanization

These figures further corroborate the notion of Java as the employment center of the country where the island hosts a wide array of industries including agriculture sector (Figure 30), industrial sector (Figure 31), financial services sector (Figure 32), trade, hotel, and restaurant sector (Figure 33), and social services sector (Figure 34).

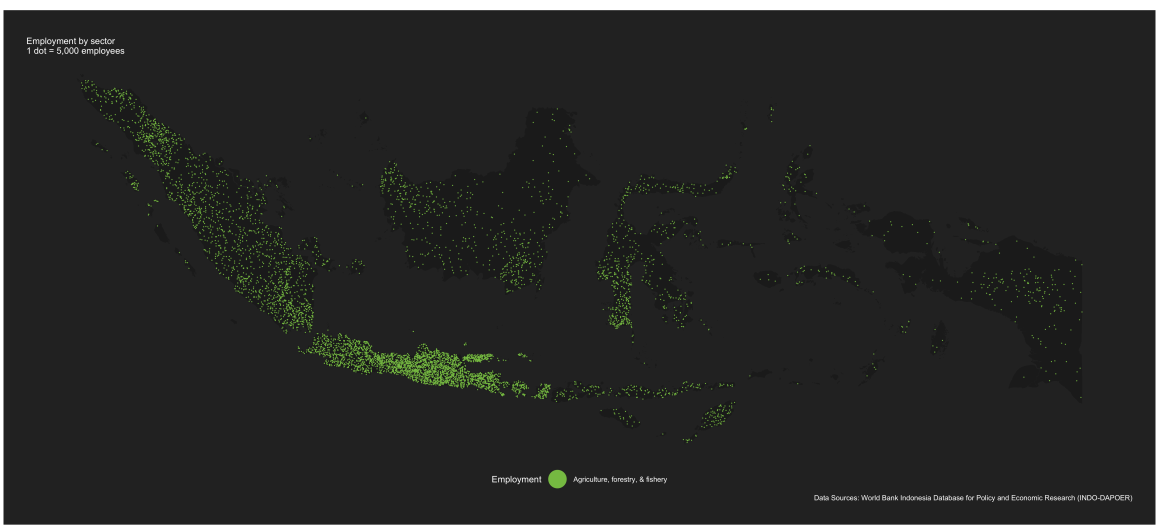

Figure 30 – Dot density map illustrating the distribution of employment in 'Agriculture' sector, 2015

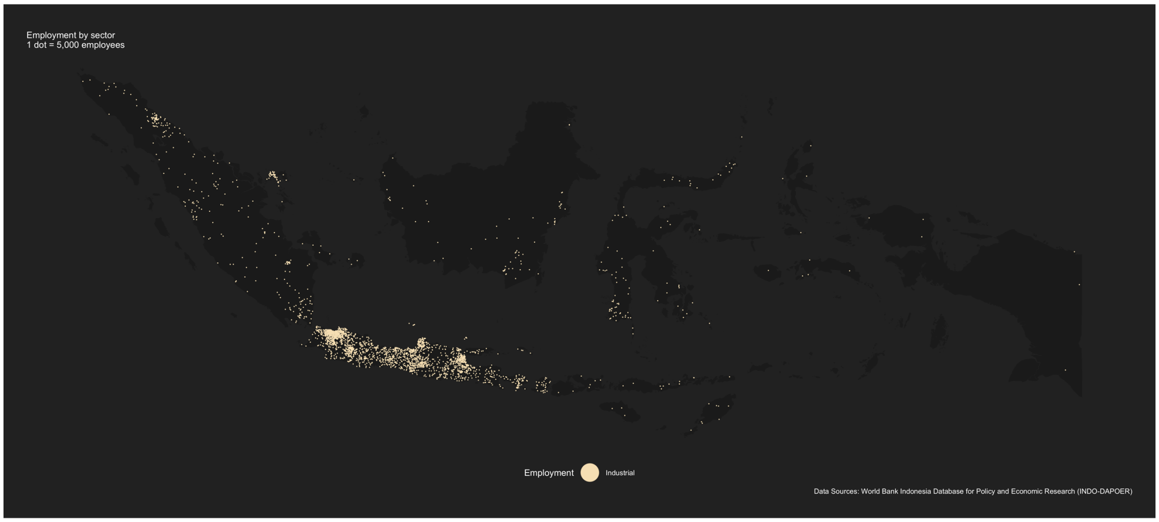

Figure 31 – Dot density map illustrating the distribution of employment in 'Industrial' sector, 2015

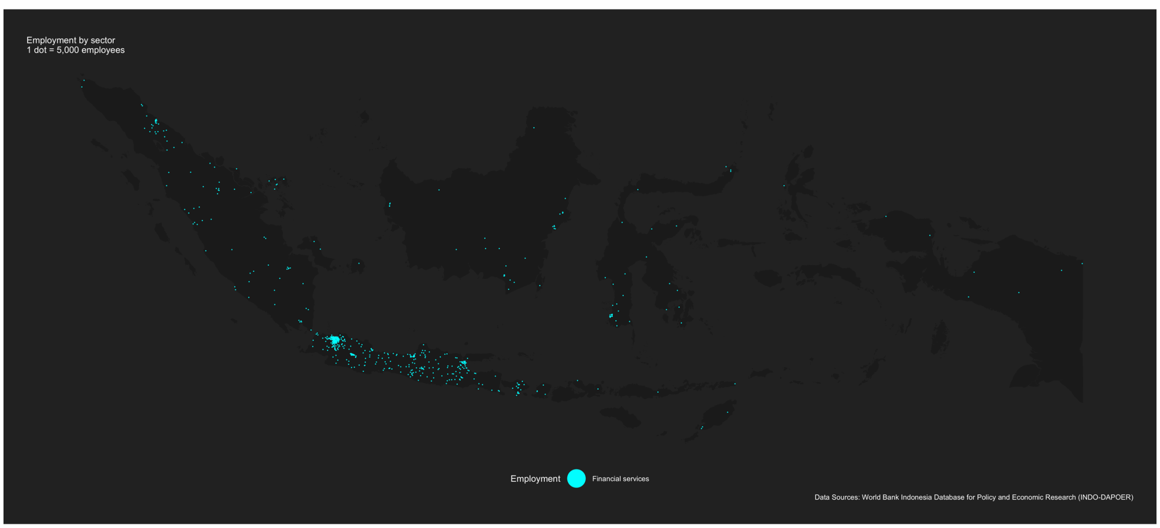

Figure 32 – Dot density map illustrating the distribution of employment in 'Financial services' sector, 2015

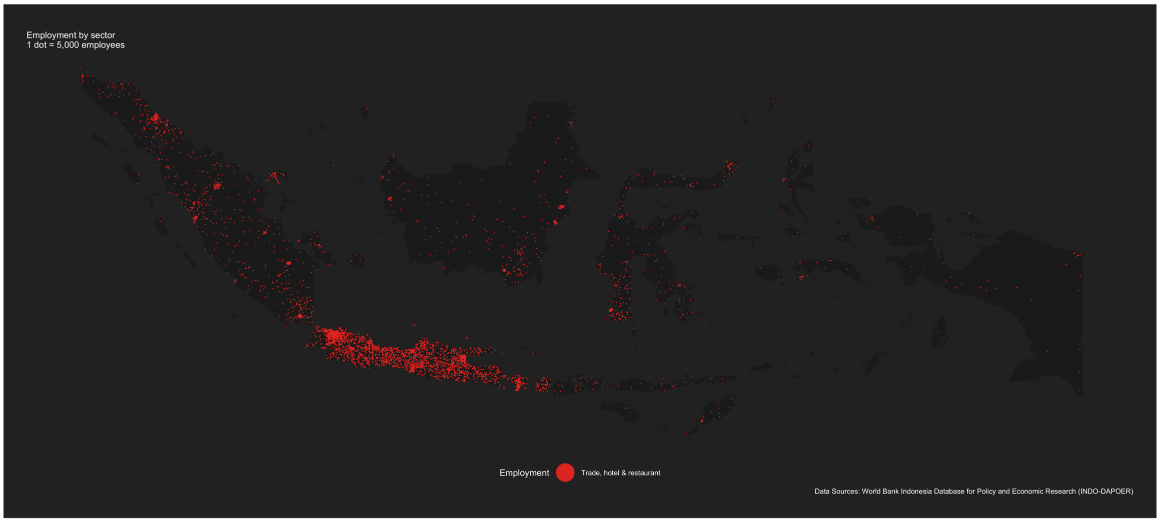

Figure 33 – Dot density map illustrating the distribution of employment in 'Trade, hotel, and restaurant' sector, 2015

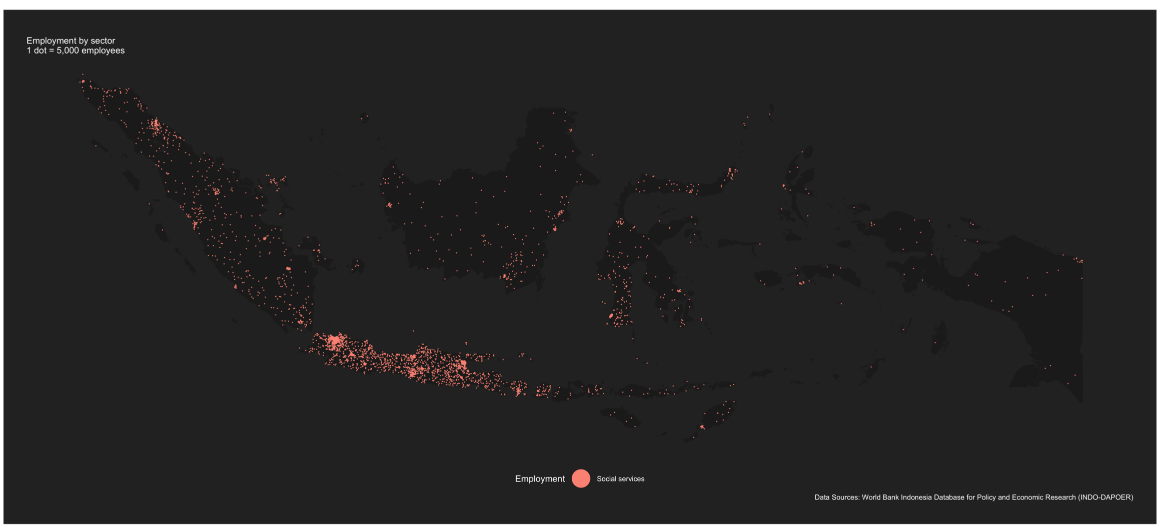

Figure 34 – Dot density map illustrating the distribution of employment in 'Social services' sector, 2015

A.1 Appendix – Chapter 5: Travel Behavior of Urban Non-Movers from a Panel Perspective

A.1.1 Cross-sectional Linear Regression Predicting Household Transportation Expenditure

This section describes the factors associated with household transportation expenditure using a cross-sectional data derived from the IFLS 4 (2007). The purpose is to initially test the hypothesis on the relationship between built environment indicators and household transportation expenditure using a linear regression model (Table 29).

Table 29 – Linear regression predicting household transportation expenditure, 2007

	Dependent variable: Monthly household transportation expenditure from a total income (logged)			
	Exclude Socioeconomic (1)		**Include Socioeconomic (2)**	
	Estimate	Std. Error	Estimate	Std. Error
Socioeconomic				
Household (HH) size			-0.011	(0.021)
# Employed household members			0.013	(0.033)
# HH members <5 years			-0.094*	(0.052)
# HH members >65 years			0.054	(0.072)
Vehicle ownership			0.074**	(0.029)
Average age of HH members			-0.002	(0.003)
Income (logged)			-0.358***	(0.037)
Built Environment				
Household density	-0.00001	(0.00002)	-0.00000	(0.00002)
Retail density	0.0002	(0.0004)	0.0002	(0.0003)
Distance to city center	0.006	0.005	0.004	0.005
Observations	1,128		1,128	
Metro dummy	Yes		Yes	
R-squared	0.009		0.100	
Adjusted R-squared	0.0003		0.086	
				*p<0.1; **p<0.05; ***p<0.01

As indicated, it appears that household transportation expenditure is largely influenced by the socioeconomic factors where all statistically significant covariates are pooled under this category. None of the built environment covariates appear to exert a statistically significant influence; nonetheless, the coefficient signs seem to reflect the expectation and findings from previous studies (Guerra, 2017; Guerra et al., 2018), particularly as it relates to the variable of household density and distance to city center. That is, respondents living in a neighborhood with higher household density might have lower household transportation expenses than those who reside in a lower density neighborhood, holding other variables constant. Moreover, as respondents live farther away from city center, their transportation expenditure might increase as well, all else equal.

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Althoff, T., Sosič, R., Hicks, J. L., King, A. C., Delp, S. L., & Leskovec, J. (2017). Large-scale physical activity data reveal worldwide activity inequality. *Nature*, *547*(7663), 336–339. https://doi.org/10.1038/nature23018

Anowar, S., Eluru, N., & Miranda-Moreno, L. F. (2014). Alternative Modeling Approaches Used for Examining Automobile Ownership: A Comprehensive Review. *Transport Reviews*, *34*(4), 441–473. https://doi.org/10.1080/01441647.2014.915440

Arel-Bundock, V. (2019). *World Bank Development indicators for R* (Version 2.6.0) [Computer software].

Asad, F. H. A. (2013). *City Centres: Understanding the Travel Behaviour of Residents and the Implications for Sustainable Travel* [Phd, University of Salford]. http://usir.salford.ac.uk/29343/

Asian Development Bank. (2012). *Key Indicators for Asia and the Pacific 2012*. Asian Development Bank. https://www.adb.org//publications/key-indicators-asia-and-pacific-2012

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, *28*(25), 3083–3107. https://doi.org/10.1002/sim.3697

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Peraturan Kepala Badan Pusat Statistik Nomor 37 Tahun 2010 tentang Klasifikasi Perkotaan dan Perdesaan di Indonesia, 37 (2010). https://jdih.bps.go.id/files/produk_hukum/perka/P08201037.pdf

Badan Pusat Statistik. (2017). *Perkembangan Jumlah Kendaraan Bermotor Menurut Jenis, 1949-2016*. Badan Pusat Statistik. https://www.bps.go.id/linkTableDinamis/view/id/1133

Bairoch, P. (1988). *Cities and Economic Development: From the Dawn of History to the Present* (C. Braider, Trans.). The University of Chicago Press. https://www.press.uchicago.edu/ucp/books/book/chicago/C/bo5971175.html

Baker, J. L., & Gadgil, G. U. (Eds.). (2017). *East Asia and Pacific Cities: Expanding Opportunities for the Urban Poor*. World Bank. https://openknowledge.worldbank.org/handle/10986/27614

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., & Grothendieck, G. (2019). *Package 'lme4.'*

Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand* (Vol. 9). MIT press.

Bento, A. M., Cropper, M. L., Mobarak, A. M., & Vinha, K. (2003). *The Impact of Urban Spatial Structure on Travel Demand in the United States*. World Bank. http://documents.worldbank.org/curated/en/649691468764046191/pdf/multi0page.pdf

Berke, E. M., Koepsell, T. D., Moudon, A. V., Hoskins, R. E., & Larson, E. B. (2007). Association of the Built Environment With Physical Activity and Obesity in Older Persons. *American Journal of Public Health*, *97*(3), 486–492. https://doi.org/10.2105/AJPH.2006.085837

Bhat, C. R., Sen, S., & Eluru, N. (2009). The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research Part B: Methodological*, *43*(1), 1–18. https://doi.org/10.1016/j.trb.2008.06.009

Boarnet, M. G., & Crane, R. (2001). *Travel by Design: The Influence of Urban Form on Travel* (1 edition). Oxford University Press.

Boarnet, M. G., Greenwald, M., & McMillan, T. E. (2008). Walking, Urban Design, and Health: Toward a Cost-Benefit Analysis Framework. *Journal of Planning Education and Research*, *27*(3), 341–358. https://doi.org/10.1177/0739456X07311073

Boarnet, M. G., Joh, K., Siembab, W., Fulton, W., & Nguyen, M. T. (2011). Retrofitting the Suburbs to Increase Walking: Evidence from a Land-use-Travel Study. *Urban Studies*, *48*(1), 129–159. https://doi.org/10.1177/0042098010364859

Boer, R., Zheng, Y., Overton, A., Ridgeway, G. K., & Cohen, D. A. (2007). Neighborhood Design and Walking Trips in Ten U.S. Metropolitan Areas. *American Journal of Preventive Medicine*, *32*(4), 298–304. https://doi.org/10.1016/j.amepre.2006.12.012

Boisjoly, G., Moreno-Monroy, A. I., & El-Geneidy, A. (2017). Informality and accessibility to jobs by public transit: Evidence from the São Paulo Metropolitan Region. *Journal of Transport Geography*, *64*, 89–96. https://doi.org/10.1016/j.jtrangeo.2017.08.005

Bonell, C. P., Hargreaves, J., Cousens, S., Ross, D., Hayes, R., Petticrew, M., & Kirkwood, B. R. (2011). Alternatives to randomisation in the evaluation of public health interventions: Design challenges and solutions. *Journal of Epidemiology and Community Health*, *65*(7), 582–587. https://doi.org/10.1136/jech.2008.082602

Brathwaite, T., & Walker, J. L. (2018). Causal inference in travel demand modeling (and the lack thereof). *Journal of Choice Modelling*, *26*, 1–18. https://doi.org/10.1016/j.jocm.2017.12.001

Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology*, *163*(12), 1149–1156. https://doi.org/10.1093/aje/kwj149

Buchanan, N., & Barnett, R. (2006). Peripheral Residential Relocation and Travel Pattern Change. *Urban Policy and Research*, *24*(2), 217–236. https://doi.org/10.1080/08111140600703824

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31–72.

Cao, X., Handy, S. L., & Mokhtarian, P. L. (2006). The Influences of the Built Environment and Residential Self-Selection on Pedestrian Behavior: Evidence from Austin, TX. *Transportation*, *33*(1), 1–20. https://doi.org/10.1007/s11116-005-7027-2

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2009a). Examining the Impacts of Residential Self-Selection on Travel Behaviour: A Focus on Empirical Findings. *Transport Reviews*, *29*(3), 359–395. https://doi.org/10.1080/01441640802539195

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2009b). The relationship between the built environment and nonwork travel: A case study of Northern California. *Transportation Research Part A: Policy and Practice*, *43*(5), 548–559. https://doi.org/10.1016/j.tra.2009.02.001

Cervero, R. (2002). Built environments and mode choice: Toward a normative framework. *Transportation Research Part D: Transport and Environment*, *7*(4), 265–284. https://doi.org/10.1016/S1361-9209(01)00024-4

Cervero, R., & Duncan, M. (2006). 'Which Reduces Vehicle Travel More: Jobs-Housing Balance or Retail-Housing Mixing? *Journal of the American Planning Association*, *72*(4), 475–490. https://doi.org/10.1080/01944360608976767

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

Chatman, D. G. (2009). Residential Choice, the Built Environment, and Nonwork Travel: Evidence Using New Data and Methods. *Environment and Planning A*, *41*(5), 1072–1089. https://doi.org/10.1068/a4114

Chatterjee, K., & Scheiner, J. (2015, July 19). *Understanding changing travel behaviour over the life course: Contributions from biographical research*. 14th International Conference on Travel Behaviour Research,14th International Conference on Travel Behaviour Research, Windsor, UK. http://eprints.uwe.ac.uk/28177/10/Chatterjee%20and%20Scheiner%20IATBR%20resource%20paper%20version%20161015.docx

Chetty, R., & Hendren, N. (2018). The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *The Quarterly Journal of Economics*. https://doi.org/10.1093/qje/qjy007

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, *129*(4), 1553–1623. https://doi.org/10.1093/qje/qju022

Christiani, Y., Byles, J. E., Tavener, M., & Dugdale, P. (2015). Do women in major cities experience better health? A comparison of chronic conditions and their risk factors between women living in major cities and other cities in Indonesia. *Global Health Action*, *8*(1), 28540. https://doi.org/10.3402/gha.v8.28540

Clark, B., Chatterjee, K., & Melia, S. (2016). Changes in level of household car ownership: The role of life events and spatial context. *Transportation*, *43*(4), 565–599. https://doi.org/10.1007/s11116-015-9589-y

Clark, B., Chatterjee, K., Melia, S., Knies, G., & Laurie, H. (2014). Life Events and Travel Behavior: Exploring the Interrelationship Using UK Household Longitudinal Study Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2413*, 54–64. https://doi.org/10.3141/2413-06

Coevering, P. van de, Maat, C., Kroesen, M., & van Wee, G. P. (2016). Causal effects of built environment characteristics on travel behaviour: A longitudinal approach. *European Journal of Transport and Infrastructure Research*, *16*(4). http://resolver.tudelft.nl/uuid:ce512c46-f922-45b5-8982-7d09597416fd

Coevering, P. van de, Maat, K., & Wee, B. van. (2015). Multi-period Research Designs for Identifying Causal Effects of Built Environment Characteristics on Travel Behaviour. *Transport Reviews*, *35*(4), 512–532. https://doi.org/10.1080/01441647.2015.1025455

Colonescu, C. (2016). *Principles of Econometrics with R*. https://bookdown.org/ccolonescu/RPoE4/

Comendador, J., & López-Lambas, M. E. (2016). Pain and Joy of a Panel Survey on Transport Studies. *Transportation Research Procedia*, *18*, 248–255. https://doi.org/10.1016/j.trpro.2016.12.034

Croissant, Y., Millo, G., Tappe, K., Toomet, O., Kleiber, C., Zeileis, A., Henningsen, A., Andronic, L., & Schoenfelder, N. (2017). *plm: Linear Models for Panel Data* (Version 1.6-6) [Computer software]. https://CRAN.R-project.org/package=plm

Dargay, J., & Hanly, M. (2003). *A panel data exploration of travel to work*. European Transport Conference, Strasbourg, France. https://aetransport.org/public/downloads/kPecA/753-514ec50e64549.pdf

Diamond, A., & Sekhon, J. S. (2006). *Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies*. https://escholarship.org/uc/item/8gx4v5qt

Endsley, K. A. (2016, April). *Diagnostics for fixed effects panel models in R*. http://karthur.org/2016/fixed-effects-panel-models-in-r.html

Escobar, A. (2011). *Encountering Development: The Making and Unmaking of the Third World*. Princeton University Press.

European Union. (2016). *The State of European Cities 2016: Cities leading the way to a better future*. European Commission. http://ec.europa.eu/regional_policy/sources/policy/themes/cities-report/state_eu_cities2016_en.pdf

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Ewing, R., & Cervero, R. (2017). "Does Compact Development Make People Drive Less?" The Answer Is Yes. *Journal of the American Planning Association*, *83*(1), 19–25. https://doi.org/10.1080/01944363.2016.1245112

Ewing, R., Greenwald, M. J., Zhang, M., Bogaerts, M., & Greene, W. (2013). Predicting Transportation Outcomes for LEED Projects. *Journal of Planning Education and Research*, *33*(3), 265–279. https://doi.org/10.1177/0739456X13482978

Fan, Y. (2007). *The built environment, activity space, and time allocation: An activity-based framework for modeling the land use and travel connection* [Ph.D., The University of North Carolina at Chapel Hill]. https://search.proquest.com/docview/304829701/abstract/F403D7C7242D47CFPQ/1

Fatmi, M. R., & Habib, M. A. (2017). Modelling mode switch associated with the change of residential location. *Travel Behaviour and Society*, *9*, 21–28. https://doi.org/10.1016/j.tbs.2017.07.006

Frank, L., Bradley, M., Kavage, S., Chapman, J., & Lawton, T. K. (2008). Urban form, travel time, and cost relationships with tour complexity and mode choice. *Transportation*, *35*(1), 37–54. https://doi.org/10.1007/s11116-007-9136-6

Gaduh, A., Gracner, T., & Rothenberg, A. D. (2017). *Improving Mobility in Developing Country Cities: Evaluating Bus Rapid Transit and Other Policies in Jakarta*.

Garrow, L. A. (2013). *Fundamentals to Discrete Choice Models*.

Gehrke, S. R., & Welch, T. F. (2017). The built environment determinants of activity participation and walking near the workplace. *Transportation*, *44*(5), 941–956. https://doi.org/10.1007/s11116-016-9687-5

Gibson, J., & Olivia, S. (2010). The Effect of Infrastructure Access and Quality on Non-Farm Enterprises in Rural Indonesia. *World Development*, *38*(5), 717–726.

Greenwald, M., & Boarnet, M. (2001). Built Environment as Determinant of Walking Behavior: Analyzing Nonwork Pedestrian Travel in Portland, Oregon. *Transportation Research Record: Journal of the Transportation Research Board*, *1780*, 33–41. https://doi.org/10.3141/1780-05

Guerra, E. (2013). *The New Suburbs: Evolving travel behavior, the built environment, and subway investment in Mexico City* [UC Berkeley]. https://escholarship.org/uc/item/4hf3b46g

Guerra, E. (2014). The Built Environment and Car Use in Mexico City: Is the Relationship Changing over Time? *Journal of Planning Education and Research*, *34*(4), 394–408. https://doi.org/10.1177/0739456X14545170

Guerra, E. (2017). Does where you live affect how much you spend on transit? The link between urban form and household transit expenditures in Mexico City. *Journal of Transport and Land Use*, *10*(1). https://doi.org/10.5198/jtlu.2017.948

Guerra, E., Caudillo, C., Goytia, C., Quiros, T. P., & Rodriguez, C. (2018). Residential location, urban form, and household transportation spending in Greater Buenos Aires. *Journal of Transport Geography*, *72*, 76–85. https://doi.org/10.1016/j.jtrangeo.2018.08.018

Halaby, C. N. (2004). Panel models in sociological research: Theory into practice. *Annual Review of Sociology*, *30*, 507–544.

Handy, S. (2017). Thoughts on the Meaning of Mark Stevens's Meta-Analysis. *Journal of the American Planning Association*, *83*(1), 26–28. https://doi.org/10.1080/01944363.2016.1246379

Handy, S., Cao, X., & Mokhtarian, P. L. (2006). Self-Selection in the Relationship between the Built Environment and Walking: Empirical Evidence from Northern California.

Journal of the American Planning Association, *72*(1), 55–74. https://doi.org/10.1080/01944360608976724

Hanly, M., & Dargay, J. (2000). Car Ownership in Great Britain: Panel Data Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, *1718*, 83–89. https://doi.org/10.3141/1718-11

Hao, L., Wang, W., & Xie, G. (2014). Rural Panel Surveys in Developing Countries: A Selective Review. *Economic and Political Studies*, *2*(2), 151–177. https://doi.org/10.1080/20954816.2014.11673849

Heres-Del-Valle, D., & Niemeier, D. (2011). CO2 emissions: Are land-use changes enough for California to reduce VMT? Specification of a two-part model with instrumental variables. *Transportation Research Part B: Methodological*, *45*(1), 150–161. https://doi.org/10.1016/j.trb.2010.04.001

Ho, D., Imai, K., King, G., Stuart, E., & Whitworth, A. (2018). *Package 'MatchIt.'*

Holt, D., Steel, D. G., Tranmer, M., & Wrigley, N. (1996). Aggregation and Ecological Effects in Geographically Based Data. *Geographical Analysis*, *28*, 244–261.

Holtzclaw, J., Clear, R., Dittmar, H., Goldstein, D., & Haas, P. (2002). Location Efficiency: Neighborhood and Socio-Economic Characteristics Determine Auto Ownership and Use - Studies in Chicago, Los Angeles and San Francisco. *Transportation Planning and Technology*, *25*(1), 1–27. https://doi.org/10.1080/03081060290032033

Hong, J., Shen, Q., & Zhang, L. (2014). How do built-environment factors affect travel behavior? A spatial analysis at different geographic scales. *Transportation*, *41*(3), 419–440. https://doi.org/10.1007/s11116-013-9462-9

Hook, W., & Replogle, M. (1996). Motorization and non-motorized transport in Asia: Transport system evolution in China, Japan and Indonesia. *Land Use Policy*, *13*(1), 69–84. https://doi.org/10.1016/0264-8377(95)00025-9

Johansson, M. (2005). Childhood Influences on Adult Travel Mode Choice. In G. Underwood (Ed.), *Traffic and Transport Psychology: Theory and Application*. Elsevier.

Jong, G. de, Fox, J., Daly, A., Pieters, M., & Smit, R. (2004). Comparison of car ownership models. *Transport Reviews*, *24*(4), 379–408. https://doi.org/10.1080/0144164032000138733

Kahn-Lang, A., & Lang, K. (2019). The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications. *Journal of Business & Economic Statistics*, 1–14. https://doi.org/10.1080/07350015.2018.1546591

Khattak, A. J., & Rodriguez, D. (2005). Travel behavior in neo-traditional neighborhood developments: A case study in USA. *Transportation Research Part A: Policy and Practice*, *39*(6), 481–500. https://doi.org/10.1016/j.tra.2005.02.009

Kitamura, R. (1990). Panel analysis in transportation planning: An overview. *Transportation Research Part A: General*, *24*(6), 401–415. https://doi.org/10.1016/0191-2607(90)90032-2

Kitamura, R., Mokhtarian, P. L., & Laidet, L. (1997). A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. *Transportation*, *24*(2), 125–158. https://doi.org/10.1023/A:1017959825565

Klinger, T., & Lanzendorf, M. (2016). Moving between mobility cultures: What affects the travel behavior of new residents? *Transportation*, *43*(2), 243–271. https://doi.org/10.1007/s11116-014-9574-x

Klöckner, C. (2004). *How single events change travel mode choice: A life span perspective*. International Conference on Traffic and Transportation Psychology. http://hdl.handle.net/11250/2392913

Knaap, G.-J., Avin, U., & Fang, L. (2017). Driving and Compact Growth: A Careful Look in the Rearview Mirror. *Journal of the American Planning Association*, *83*(1), 32–35. https://doi.org/10.1080/01944363.2017.1251276

Kraus, M. (1977). A Theoretical Justification of the “Disutility” Model of Travel Demand. *International Journal of Transport Economics / Rivista Internazionale Di Economia Dei Trasporti*, *4*(2), 107–113.

Krizek, K. J. (2003). Residential Relocation and Changes in Urban Travel: Does Neighborhood-Scale Urban Form Matter? *Journal of the American Planning Association*, *69*(3), 265–281. https://doi.org/10.1080/01944360308978019

Kroesen, M., & Goulias, K. G. (2016). Modelling activity-travel behaviour dynamics with panel data: The state-of-the-art. *European Journal of Transport and Infrastructure Research*, *16*(4), 633–637. Scopus.

Kuh, D. J., & Cooper, C. (1992). Physical activity at 36 years: Patterns and childhood predictors in a longitudinal study. *Journal of Epidemiology and Community Health*, *46*(2), 114–119.

Kuzmyak, J., Baber, C., & Savory, D. (2006). Use of Walk Opportunities Index to Quantify Local Accessibility. *Transportation Research Record: Journal of the Transportation Research Board*, *1977*, 145–153. https://doi.org/10.3141/1977-19

Lanzendorf, M. (2003). *Mobility biographies: A new perspective for understanding travel behaviour*. 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland. http://archiv.ivt.ethz.ch/news/archive/20030810_IATBR/lanzendorf.pdf

Lave, C. A. (1969). A behavioral approach to modal split forecasting. *Transportation Research*, *3*(4), 463–480. https://doi.org/10.1016/0041-1647(69)90041-0

Lee, C., & Moudon, A. V. (2006). The 3Ds+R: Quantifying land use and urban form correlates of walking. *Transportation Research Part D: Transport and Environment*, *11*(3), 204–215. https://doi.org/10.1016/j.trd.2006.02.003

Lee, J. S., Zegras, C., Ben-Joseph, E., & Park, S. (2014). Does urban living influence baby boomers' travel behavior? *Journal of Transport Geography*, *35*, 21–29. https://doi.org/10.1016/j.jtrangeo.2014.01.004

Lee, Y., & Guhathakurta, S. (2018). An analysis of the effects of suburban densification on vehicle use for shopping: Do existing residents respond to land-use changes in the same way as recent movers? *Journal of Transport Geography*, *68*, 193–204. https://doi.org/10.1016/j.jtrangeo.2018.03.011

Li, F., Fisher, K. J., Brownson, R. C., & Bosworth, M. (2005). Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. *Journal of Epidemiology & Community Health*, *59*(7), 558–564. https://doi.org/10.1136/jech.2004.028399

Lim, S., Marcus, S. M., Singh, T. P., Harris, T. G., & Levanon Seligson, A. (2014). Bias Due to Sample Selection in Propensity Score Matching for a Supportive Housing Program Evaluation in New York City. *PLoS ONE*, *9*(10). https://doi.org/10.1371/journal.pone.0109112

Lin, T., Wang, D., & Zhou, M. (2018). Residential relocation and changes in travel behavior: What is the role of social context change? *Transportation Research Part A: Policy and Practice*, *111*, 360–374. https://doi.org/10.1016/j.tra.2018.03.015

Long, L. (1997). Panels for Transportation Planning: Theoretical Issues and Empirical Challenges. In T. F. Golob, R. Kitamura, & L. Long (Eds.), *Panels for Transportation Planning—Methods and Applications*. Springer US. https://www.springer.com/us/book/9780792399667

Lu, Y. (2010). Rural-urban migration and health: Evidence from longitudinal data in Indonesia. *Social Science & Medicine*, *70*(3), 412–419. https://doi.org/10.1016/j.socscimed.2009.10.028

Macfarlane, G. S., Garrow, L. A., & Mokhtarian, P. L. (2015). The influences of past and present residential locations on vehicle ownership decisions. *Transportation Research Part A: Policy and Practice*, *74*, 186–200. https://doi.org/10.1016/j.tra.2015.01.005

Majid, M. R., Nordin, A. N., & Medugu, I. N. (2014). Influence of housing development designs on household vehicle miles traveled: A case of Iskandar Malaysia. *Transportation Research Part D: Transport and Environment*, *33*, 63–73. https://doi.org/10.1016/j.trd.2014.09.001

Manville, M. (2017). Travel and the Built Environment: Time for Change. *Journal of the American Planning Association*, *83*(1), 29–32. https://doi.org/10.1080/01944363.2016.1249508

McKenzie, D. (2020, January 21). *Revisiting the Difference-in-Differences Parallel Trends Assumption: Part I Pre-Trend Testing*. https://blogs.worldbank.org/impactevaluations/revisiting-difference-differences-parallel-trends-assumption-part-i-pre-trend

Mjahed, L. B., Frei, C., & Mahmassani, H. S. (2015). Walking Behavior: The Role of Childhood Travel Experience. *Transportation Research Record: Journal of the Transportation Research Board*, *2495*, 94–100. https://doi.org/10.3141/2495-10

Monkkonen, P. (2013). Urban land-use regulations and housing markets in developing countries: Evidence from Indonesia on the importance of enforcement. *Land Use Policy*, *34*, 255–264. https://doi.org/10.1016/j.landusepol.2013.03.015

Müggenburg, H., Busch-Geertsema, A., & Lanzendorf, M. (2015). Mobility biographies: A review of achievements and challenges of the mobility biographies approach and a framework for further research. *Journal of Transport Geography*, *46*, 151–163. https://doi.org/10.1016/j.jtrangeo.2015.06.004

Næss, P., Peters, S., Stefansdottir, H., & Strand, A. (2018). Causality, not just correlation: Residential location, transport rationales and travel behavior across metropolitan contexts. *Journal of Transport Geography*, *69*, 181–195. https://doi.org/10.1016/j.jtrangeo.2018.04.003

Nasri, A., Carrion, C., Zhang, L., & Baghaei, B. (2018). Using propensity score matching technique to address self-selection in transit-oriented development (TOD) areas. *Transportation*. https://doi.org/10.1007/s11116-018-9887-2

Nasri, A., & Zhang, L. (2012). Impact of Metropolitan-Level Built Environment on Travel Behavior. *Transportation Research Record*, *2323*(1), 75–79. https://doi.org/10.3141/2323-09

Olmos, A., & Govindasamy, P. (2015). Propensity Scores: A Practical Introduction Using R. *Journal of MultiDisciplinary Evaluation*, *11*(25), 68–88.

Paaswell, R. E. (1997). Why Panels for Transportation Planning? In T. F. Golob, R. Kitamura, & L. Long (Eds.), *Panels for Transportation Planning—Methods and Applications* (pp. 4–14). Springer US. https://www.springer.com/us/book/9780792399667

Parady, G. T., Chikaraishi, M., Takami, K., & Harata, N. (2014). A Panel Data Approach to Understanding the Effect of the Built Environment on Travel Behavior. *Urban and Regional Planning Review*, *1*, 18–38. https://doi.org/10.14398/urpr.1.18

Park, K., Ewing, R., Scheer, B. C., & Ara Khan, S. S. (2018). Travel Behavior in TODs vs. Non-TODs: Using Cluster Analysis and Propensity Score Matching. *Transportation Research Record*, 0361198118774159. https://doi.org/10.1177/0361198118774159

Pas, E. I. (1988). Weekly travel-activity behavior. *Transportation*, *15*(1), 89–109. https://doi.org/10.1007/BF00167982

Pendyala, R. M., & Bhat, C. R. (2017). *Emerging Issues in Travel Behavior Analysis*. http://www.caee.utexas.edu/prof/bhat/ABSTRACTS/Workshop_EmergingIssues_Full.pdf

Prillwitz, J., Harms, S., & Lanzendorf, M. (2006). Impact of Life-Course Events on Car Ownership. *Transportation Research Record*, *1985*(1), 71–77. https://doi.org/10.1177/0361198106198500108

Prillwitz, J., Harms, S., & Lanzendorf, M. (2007). Interactions between Residential Relocations, Life Course Events, and Daily Commute Distances. *Transportation Research Record*, *2021*(1), 64–69. https://doi.org/10.3141/2021-08

Purvis, C. L. (1997). Planning for Panel Surveys in the San Francisco Bay Area. In T. F. Golob, R. Kitamura, & L. Long (Eds.), *Panels for Transportation Planning—Methods and Applications* (pp. 193–203). Springer US. https://www.springer.com/us/book/9780792399667

Qibthiyyah, R., & Utomo, A. J. (2016). Family Matters: Demographic Change and Social Spending in Indonesia. *Bulletin of Indonesian Economic Studies*, *52*(2), 133–159. https://doi.org/10.1080/00074918.2016.1211077

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Raimond, T., & Hensher, D. A. (1997). A Review of Empirical Studies and Applications. In T. F. Golob, R. Kitamura, & L. Long (Eds.), *Panels for Transportation Planning—Methods and Applications* (pp. 15–72). Springer US. https://www.springer.com/us/book/9780792399667

Ralph, K. M. (2018). Childhood Car Access: Long-term Consequences for Education, Employment, and Earnings. *Journal of Planning Education and Research*, 0739456X18798451. https://doi.org/10.1177/0739456X18798451

Ravallion, M. (2008). Evaluating Anti-Poverty Programs. In *Handbook of Development Economics* (pp. 3787–3846). Elsevier. https://econpapers.repec.org/bookchap/eeedevchp/5-59.htm

Robinson, W. S. (1950). Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, *15*(3), 351–357. https://doi.org/10.2307/2087176

Rodríguez, D. A., & Joo, J. (2004). The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment*, *9*(2), 151–173. https://doi.org/10.1016/j.trd.2003.11.001

Rosales-Rueda, M., & Triyana, M. (2018). The Persistent Effects of Early-Life Exposure to Air Pollution: Evidence from the Indonesian Forest Fires. *Journal of Human Resources*, 0117-8497R1. https://doi.org/10.3368/jhr.54.4.0117.8497R1

Rubin, J. (2011). *Choosing Transit: The Influence of Past Travel Behavior, Attitudes and Habits on Present Choices* [UC Berkeley]. https://escholarship.org/uc/item/1vs4846h

Rumbach, A., & Shirgaokar, M. (2017). Predictors of household exposure to monsoon rain hazards in informal settlements. *Natural Hazards*, *85*(2), 709–728. https://doi.org/10.1007/s11069-016-2599-z

Salomon, I., & Ben-Akiva, M. (1983). The Use of the Life-Style Concept in Travel Demand Models. *Environment and Planning A: Economy and Space*, *15*(5), 623–638. https://doi.org/10.1068/a150623

Scheiner, J. (2007). Mobility Biographies: Elements of a Biographical Theory of Travel Demand. *Erdkunde*, *61*(2), 161–173.

Scheiner, J., & Holz-Rau, C. (2013a). A comprehensive study of life course, cohort, and period effects on changes in travel mode use. *Transportation Research Part A: Policy and Practice*, *47*, 167–181. https://doi.org/10.1016/j.tra.2012.10.019

Scheiner, J., & Holz-Rau, C. (2013b). Changes in travel mode use after residential relocation: A contribution to mobility biographies. *Transportation*, *40*(2), 431–458. https://doi.org/10.1007/s11116-012-9417-6

Setiawan, B. (2010). *Kampung Kota Dan Kota Kampung: Tantangan Perencanaan Kota di Indonesia*. Universitas Gadjah Mada.

Sheng, Y. K. (2012). The Challenges of Promoting Productive, Inclusive, and Sustainable Urbanization. In Y. K. Sheng & M. Thuzar (Eds.), *Urbanization in Southeast Asia: Issues and Impacts* (pp. 10–69). Institute of Southeast Asian Studies.

Small, K. A., & Winston, C. W. (1999). The Demand for Transportation: Model and Applications. In J. A. Gomez-Ibanez, W. B. Tye, & C. W. Winston (Eds.), *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer*. Brookings Institution Press.

Smart, M. J., & Klein, N. J. (2018a). Complicating the Story of Location Affordability. *Housing Policy Debate*, *28*(3), 393–410. https://doi.org/10.1080/10511482.2017.1371784

Smart, M. J., & Klein, N. J. (2018b). Remembrance of Cars and Buses Past: How Prior Life Experiences Influence Travel. *Journal of Planning Education and Research*, *38*(2), 139–151. https://doi.org/10.1177/0739456X17695774

Spears, S., Boarnet, M. G., & Houston, D. (2016). Driving reduction after the introduction of light rail transit: Evidence from an experimental-control group evaluation of the Los Angeles Expo Line. *Urban Studies*, 0042098016657261. https://doi.org/10.1177/0042098016657261

Stanbridge, K., & Lyons, G. (2006, August). *Travel behaviour considerations during the process of residential relocation*. 11th International Conference on Travel Behaviour Research,11th International Conference on Travel Behaviour Research, Kyoto, Japan. http://eprints.uwe.ac.uk/9789/

Starks, H., & Garrido, M. (2014, October 20). *Observational & Quasi-experimental Research Methods*. 8th Annual Kathleen Foley Palliative Care Retreat Methods Workshop. http://www.npcrc.org/files/NPCRC.Observational-PropensityScoreMethodsWkshop.10-20-14.pdf

Stevens, M. R. (2017). Does Compact Development Make People Drive Less? *Journal of the American Planning Association*, *83*(1), 7–18. https://doi.org/10.1080/01944363.2016.1240044

Strauss, J., Witoelar, F., & Sikoki, B. (2016). *The Fifth Wave of the Indonesia Family Life Survey (IFLS): Overview and Field Report*. RAND. https://www.rand.org/pubs/working_papers/WR1143z1.html

Strauss, J., Witoelar, F., Sikoki, B., & Wattie, A. M. (2009). *The Fourth Wave of the Indonesia Family Life Survey (IFLS): Overview and Field Report*. RAND. https://www.rand.org/labor/FLS/IFLS/download.html

Stuart, E. A., Cole, S. R., Bradshaw, C. P., & Leaf, P. J. (2001). The use of propensity scores to assess the generalizability of results from randomized trials. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)*, *174*(2), 369–386. https://doi.org/10.1111/j.1467-985X.2010.00673.x

Sujarwoto, S., Tampubolon, G., & Pierewan, A. C. (2017). Individual and Contextual Factors of Happiness and Life Satisfaction in a Low Middle Income Country. *Applied Research in Quality of Life*. https://doi.org/10.1007/s11482-017-9567-y

Swait, J., Adamowicz, W., & Bueren, M. van. (2004). Choice and temporal welfare impacts: Incorporating history into discrete choice models. *Journal of Environmental Economics and Management*, *47*(1), 94–116. https://doi.org/10.1016/S0095-0696(03)00077-9

Thigpen, C. G. (2017). *The Reciprocal Relationship between Children and Young Adults' Travel Behavior and Their Travel Attitudes, Skills, and Norms* [UC Davis].

https://ncst.ucdavis.edu/wp-content/uploads/2016/11/Thigpen_Thesis_Full_FINAL.pdf

Thomas, D., Frankenberg, E., & Smith, J. P. (2001). Lost but Not Forgotten: Attrition and Follow-up in the Indonesia Family Life Survey. *The Journal of Human Resources*, *36*(3), 556–592. https://doi.org/10.2307/3069630

Thomas, D., Witoelar, F., Frankenberg, E., Sikoki, B., Strauss, J., Sumantri, C., & Suriastini, W. (2012). Cutting the costs of attrition: Results from the Indonesia Family Life Survey. *Journal of Development Economics*, *98*(1), 108–123. https://doi.org/10.1016/j.jdeveco.2010.08.015

Tourangeau, R., Zimowski, M., & Ghadialy, R. (1997). *An introduction to panel surveys in transportation studies*. United States. Federal Highway Administration.

Underwood, S. K., & Handy, S. (2012). Adolescent Attitudes Towards Active Transportation: Bicycling in Youth in Retrospect from Adulthood. *University of California, Davis. Institute of Transportation Studies. Research Report*. https://trid.trb.org/view/1217348

Underwood, S. K., Handy, S. L., Paterniti, D. A., & Lee, A. E. (2014). Why do teens abandon bicycling? A retrospective look at attitudes and behaviors. *Journal of Transport & Health*, *1*(1), 17–24. https://doi.org/10.1016/j.jth.2013.12.002

United Nations. (2012). *World Urbanization Prospects: The 2011 Revision*. United Nations. http://esa.un.org/undp/wup/index.html

Van Wissen, L. J. G., & Meurs, H. J. (1989). The Dutch mobility panel: Experiences and evaluation. *Transportation*, *16*(2), 99–119. https://doi.org/10.1007/BF00163111

Villanueva, K., Knuiman, M., Nathan, A., Giles-Corti, B., Christian, H., Foster, S., & Bull, F. (2014). The impact of neighborhood walkability on walking: Does it differ across adult life stage and does neighborhood buffer size matter? *Health & Place*, *25*, 43–46. https://doi.org/10.1016/j.healthplace.2013.10.005

Wai-Poi, M., Alatas, H., Chandrashekar, K., & Lain, J. (2018). *The Different Faces of Urban Indonesia: Recent Urban Trends in Indonesia*.

Wang, D., & Cao, X. (2017). Impacts of the built environment on activity-travel behavior: Are there differences between public and private housing residents in Hong Kong? *Transportation Research Part A: Policy and Practice*, *103*(Supplement C), 25–35. https://doi.org/10.1016/j.tra.2017.05.018

Wang, D., & Lin, T. (2017). Built environment, travel behavior, and residential self-selection: A study based on panel data from Beijing, China. *Transportation*. https://doi.org/10.1007/s11116-017-9783-1

Wasfi, R., Steinmetz-Wood, M., & Kestens, Y. (2017). Place matters: A longitudinal analysis measuring the association between neighbourhood walkability and walking by age group and population center size in Canada. *PLOS ONE*, *12*(12), e0189472. https://doi.org/10.1371/journal.pone.0189472

Weinberger, R., & Goetzke, F. (2010). Unpacking Preference: How Previous Experience Affects Auto Ownership in the United States. *Urban Studies*, *47*(10), 2111–2128. https://doi.org/10.1177/0042098009357354

Witoelar, F. (2017). *Indonesia Family Life Survey 5 and IFLS East*. http://www.smeru.or.id/sites/default/files/events/d1_opening_ifls_firman_witoelar_0.pdf

World Bank. (2015). *East Asia's Changing Urban Landscape: Measuring a Decade of Spatial Growth*. World Bank. https://www.worldbank.org/content/dam/Worldbank/Publications/Urban%20Development/EAP_Urban_Expansion_full_report_web.pdf

World Bank. (2016). *Indonesia's Urban Story*. The World Bank. http://www.worldbank.org/en/news/feature/2016/06/14/indonesia-urban-story

World Bank. (2018a). *Indonesia Economic Quarterly: Urbanization for All*. World Bank. http://documents.worldbank.org/curated/en/545961539697659383/Indonesia-Economic-Quarterly-Urbanization-for-All

World Bank. (2018b). *World Development Indicators*. The World Bank. https://datacatalog.worldbank.org/dataset/world-development-indicators

Yamamoto, T. (2008). The Impact of Life-Course Events on Vehicle Ownership Dynamics: The Cases of France and Japan. *IATSS Research*, *32*(2), 34–43. https://doi.org/10.1016/S0386-1112(14)60207-7

Yang, J. (2006). Transportation Implications of Land Development in a Transitional Economy: Evidence from Housing Relocation in Beijing. *Transportation Research Record: Journal of the Transportation Research Board*, *1954*, 7–14. https://doi.org/10.3141/1954-02

Zahabi, S. A. H., Miranda-Moreno, L., Patterson, Z., & Barla, P. (2015). Spatio-temporal analysis of car distance, greenhouse gases and the effect of built environment: A latent class regression analysis. *Transportation Research Part A: Policy and Practice*, *77*, 1–13. https://doi.org/10.1016/j.tra.2015.04.002

Zegras, C. (2004). Influence of Land Use on Travel Behavior in Santiago, Chile. *Transportation Research Record: Journal of the Transportation Research Board*, *1898*, 175–182. https://doi.org/10.3141/1898-21

Zegras, C. (2010). The Built Environment and Motor Vehicle Ownership and Use: Evidence from Santiago de Chile. *Urban Studies*, *47*(8), 1793–1817. https://doi.org/10.1177/0042098009356125

Zhang, M. (2004). The Role of Land Use in Travel Mode Choice: Evidence from Boston and Hong Kong. *Journal of the American Planning Association*, *70*(3), 344–360. https://doi.org/10.1080/01944360408976383

Zhang, M., & Kukadia, N. (2005). Metrics of Urban Form and the Modifiable Areal Unit Problem. *Transportation Research Record: Journal of the Transportation Research Board*, *1902*, 71–79. https://doi.org/10.3141/1902-09

Zhang, M., & Zhang, W. (2018). When Context Meets Self-Selection: The Built Environment–Travel Connection Revisited. *Journal of Planning Education and Research*, 0739456X18755495. https://doi.org/10.1177/0739456X18755495

Zhou, B., & Kockelman, K. (2008). Self-Selection in Home Choice: Use of Treatment Effects in Evaluating Relationship Between Built Environment and Travel Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2077*, 54–61. https://doi.org/10.3141/2077-08

Zumkeller, D., & Chlond, B. (2009). *Dynamics of Change: Fifteen-Year German Mobility Panel*. Transportation Research Board 88th Annual MeetingTransportation Research Board. https://trid.trb.org/view/880678